UNIVERSITY
OF MICHIGAN
OCT 20 1950
MATHEMATICS
LIBRARY

PROCEEDINGS

OF THE

AMERICAN MATHEMATICAL SOCIETY

D
c. X

EDITED BY

GUSTAV A. HEDLUND NATHAN JACOBSON

A. C. SCHAEFFER

WITH THE ASSISTANCE OF

C. B. ALLENDOERFER
M. M. DAY
G. P. HOCHSCHILD
R. H. BING
J. B. DIAZ
I. M. SHEFFER
G. W. WHITEHEAD
A. H. CLIFFORD
F. G. DRESSEL
MAX SHIFFMAN

VOLUME 1, NUMBER 4
AUGUST, 1950

PUBLISHED BY THE SOCIETY
MENASHA, WIS., AND NEW YORK

Journals Published by the

American Mathematical Society

Bulletin of the American Mathematical Society

This journal is the official organ of the Society. It reports official acts of the Society and the details of its meetings, with abstracts of all research papers presented. It contains some of the officially invited addresses presented before the Society. It contains also announcements and notes on matters of current interest, and reviews of advanced mathematical books.

One volume of six numbers is planned for 1950. The subscription price for the BULLETIN and the PROCEEDINGS (see below) is now $11.00, and the two journals are provided together to subscribers, and cannot be obtained separately. For 1951 subscriptions may be placed for the BULLETIN separately and the subscription price will be $7.00. It is planned to publish one volume of six issues in 1951. Sets of volumes of the BULLETIN are available as follows: vols. 1–36 (1894–1930), $7.00 each; vols. 37–53 (1931–1947), $9.00 each; vols. 54–55 (1948–1949), $11.00 each.

Proceedings of the American Mathematical Society

This journal is devoted entirely to research in pure and applied mathematics, and is devoted principally to the publication of original papers of moderate length. It contains the material formerly published in the even-numbered issues of the BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY. Volume 1 appears in 1950.

One volume of six numbers is planned for 1950. The PROCEEDINGS is supplied with the BULLETIN to all subscribers. See the announcement concerning the BULLETIN above. For 1951 one volume of six issues is planned and subscriptions may be placed for the PROCEEDINGS separately. The price will be $11.00.

Transactions of the American Mathematical Society

This journal is devoted entirely to research in pure and applied mathematics, and includes in general longer memoirs than the PROCEEDINGS.

Normally, two volumes of three numbers each are published annually. The price per volume is $6.00. Volumes 1–36 (1900–1934) consist of four numbers each. Volumes 66–67 (1949) consist of two numbers each. Sets of volumes are available as follows: vols. 1–21 (1900–1920), $7.00 each; vols. 22–32 (1921–1930), $8.00 each; vols. 33–36 (1931–1934), $9.00 each; vols. 37–68 (1935–1950), $6.00 each.

Mathematical Reviews

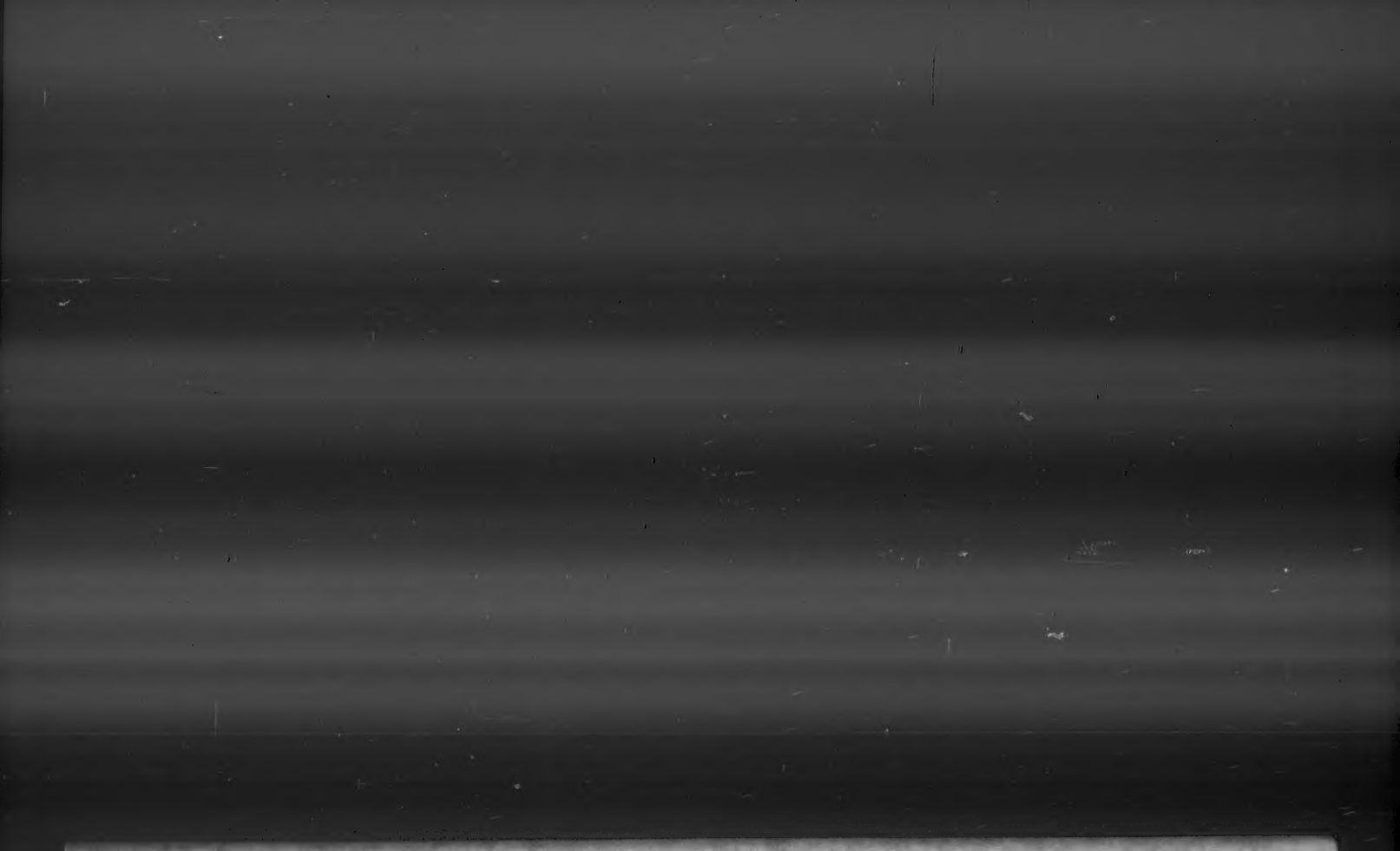
This journal contains abstracts and reviews of the current mathematical literature of the world. It is sponsored by thirteen mathematical organizations, located both in the United States and abroad.

The publication of MATHEMATICAL REVIEWS was begun in 1940. It appears monthly, excepting July, and the price per annual volume is $20.00. Orders for complete volumes only are accepted. Volumes 1–10 (1940–1949) are available at $20.00 per volume.

The PROCEEDINGS OF THE AMERICAN MATHEMATICAL SOCIETY is published bimonthly, in February, April, June, August, October, and December. Price per annual volume, with the BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY, $11.00. Subscriptions, orders for back numbers, and inquiries in regard to nondelivery of current numbers should be addressed to the American Mathematical Society, 450 Ahnaip Street, Menasha, Wis., or Low Memorial Library, 531 West 116th Street, New York 27, N.Y.

Entered as second-class matter at the post office at Menasha, Wisconsin. Authorization is granted under the authority of the act of August 24, 1912, as amended by the act of August 4, 1947 (sec. 34.21, P. L. & R.).

Copyright, American Mathematical Society, 1950



ON CHERN'S INVARIANT FOR RIEMANNIAN 4-MANIFOLDS

HANS SAMELSON

1. Introduction. S. S. Chern has defined, guided by geometric considerations, a differential geometric invariant for closed, orientable, 4-dimensional Riemannian manifolds [1];[1] the invariant is closely connected with the fact that the orthogonal group in 4 variables is essentially, that is, up to a finite covering, the direct product of two orthogonal groups in 3 variables. It is the purpose of this note to exhibit a space for which this invariant is not zero; at the same time we shall verify the well known generalized Gauss-Bonnet formula for the curvatura integra. The space is the complex projective plane K^2, in its elliptic metric; this metric is discussed in detail in [2], and we shall freely refer to this paper for the necessary formulae of Riemannian geometry in Cartan's formulation; for the basic definitions of differential forms we refer to [3]. We make use of the well known fact that K^2 is the base space of a fiber decomposition of S^5 with S^1 as fiber (see [4] and [5]).

2. The space K^2. We now consider the complex plane K^2. As is well known, it is obtained from the 5-sphere S^5 by a fiber decomposition [4]: We represent S^5 as unit sphere $\sum|z_i|^2=1$ in a complex-euclidean 3-dimensional space C^3 with coordinates (z_1, z_2, z_3). Two points y and z of S^5 are called equivalent, under the relation Θ, if there exists a real θ such that $y_k=z_k\cdot e^{i\theta}$; the equivalence classes, or fibers, are great circles, with θ as arc length.

One finds that the decomposition space S^5/Θ with the induced topology is a 4-dimensional manifold; it is called K^2, and with the canonical map ϕ: $S^5\to K^2$ the system (S^5, S^1, K^2, ϕ) is a fiber bundle [5] (with the rotations of S^1 as structure group). In addition there is a natural way of making K^2 into a differentiable (even analytic) manifold, such that ϕ is a differentiable (analytic) map: if $z^0=(z_1^0, z_2^0, z_3^0)$ with, say, $z_1^0\neq 0$, is a point of S^5, then the real and imaginary part of z_2/z_1 and z_3/z_1 serve as analytic coordinates in the neighborhood of $\phi(z^0)$. Now S^5 is a Riemannian manifold, as subspace of C^3, with fundamental form $\sum|dz_i|^2$. One verifies after some computation that for a vector $\xi=(dz_1, dz_2, dz_3)$, tangent to S^5 at (z_1, z_2, z_3), the expression $\sum|dz_i|^2-|\sum z_i\overline{dz_i}|^2$ depends only on the image vector of ξ under the map ϕ, and that it represents a positive definite

Received by the editors June 1, 1949.

[1] Numbers in brackets refer to the references cited at the end of the paper.

(analytic) quadratic form on the tangent space to K^2; the Riemannian geometry determined by it is called the elliptic hermitian geometry of K^2; from now on we consider K^2 with this metric (see [2]). Any vector of S^5, which is perpendicular to the fiber through its point of support, is mapped by ϕ into a vector of the same length. Using this fact, and considering the volume forms $(g^{1/2}dx_1 \cdots dx_n)$ of the Riemannian manifolds S^5, K^2, and S^1, one can show that the volumes $v(S^5)$, $v(K^2)$, and $v(S^1)$ satisfy the relation $v(S^5) = v(K^2) \cdot v(S^1)$; since $v(S^1) = 2\pi$, and $v(S^5) = \pi^3$ [6, p. 303], we have

$$(*) \qquad v(K^2) = \pi^2/2.$$

3. **The differential forms.** Chern's invariant and the total curvature form are expressed in terms of Cartan's differential forms [1]. We first consider the complex curvature forms Ω_{ij} for K^2; we take the necessary formulae from [2]: there are two complex 1-forms ω_1 and and ω_2, and the Ω_{ij}'s are given by

$$\Omega_{ii} = \omega_i\bar{\omega}_i + \sum \omega_k\bar{\omega}_k; \qquad \Omega_{ij} = \omega_j\bar{\omega}_i, \qquad i \neq j.$$

We have to set $\omega_1 = \theta_1 + i\psi_1$, and $\omega_2 = \theta_2 + i\psi_2$.

The volume form dv is then given by $\theta_1\theta_2\psi_1\psi_2$, and we note, in consequence of (*), §2, that

$$(**) \qquad \int_{K^2} \theta_1\theta_2\psi_1\psi_2 = \int_{K^2} dv = \frac{\pi^2}{2}.$$

Substituting into the Ω_{ij}'s we have

$$\begin{aligned}\Omega_{ii} &= (\theta_i + i\psi_i)(\theta_i - i\psi_i) + \sum (\theta_k + i\psi_k)(\theta_k - i\psi_k) \\ &= 2i\psi_i\theta_i + 2i\sum \psi_k\theta_k, \\ \Omega_{ij} &= (\theta_j + i\psi_j)(\theta_i - i\psi_i) = \theta_j\theta_i + \psi_j\psi_i + i(\psi_j\theta_i - \theta_j\psi_i), \qquad i \neq j.\end{aligned}$$

Setting $\Omega_{ij} = \Theta_{ij} + i\Psi_{ij}$, and denoting the real curvature forms by $\Omega'_{\alpha\beta}$, $\alpha, \beta = 0, 1, 2, 3$, we obtain for the matrix $\psi^r = (\Omega'_{\alpha\beta})$ by [2, (44)] the following:

$$\Omega^r = \begin{pmatrix} \Theta & -\Psi \\ \Psi & \Theta \end{pmatrix} \quad \text{with} \quad \Theta = \begin{pmatrix} 0 & \theta_2\theta_1 + \psi_2\psi_1 \\ \theta_1\theta_2 + \psi_1\psi_2 & 0 \end{pmatrix}$$

and

$$\Psi = \begin{pmatrix} 4\psi_1\theta_1 + 2\psi_2\theta_2 & \psi_1\theta_2 - \theta_1\psi_2 \\ \psi_1\theta_2 - \theta_1\psi_2 & 4\psi_2\theta_2 + 2\psi_1\theta_1 \end{pmatrix}.$$

With the quantities $\Omega'_{\alpha\beta}$ we now compute the Chern form Γ, and

the total curvature form K; they are given by $\Gamma=\sum_{\alpha<\beta}[\Omega'_{\alpha\beta}]^2$ and $K=\Omega'_{01}\Omega'_{23}+\Omega'_{02}\Omega'_{31}+\Omega'_{03}\Omega'_{12}$ (see [1, p. 970]). By a straightforward multiplication we obtain

$$\Gamma = -24\theta_1\theta_2\psi_1\psi_2 = -24dv \quad \text{and} \quad K = 24\theta_1\theta_2\psi_1\psi_2 = 24dv.$$

Incidentally this shows that Γ really is a form of K^2 and not only of a fiber bundle over K^2.

We obtain now for the characteristic, using (**),

$$\chi^{(K^2)}=\frac{1}{4\pi^2}\cdot\int_{K^2}K=\frac{6}{\pi^2}\cdot\int_{K^2}dv=\frac{6}{\pi^2}\cdot\frac{\pi^2}{2}=3,$$

which checks with well known facts, and for the Chern invariant

$$\eta=\frac{1}{2\pi^2}\cdot\int_{K^2}\Gamma=-\frac{12}{\pi^2}\cdot\int_{K^2}dv=-6,$$

which is different from 0; as predicted by the theory, the value of η is integral. The two invariants I_r and I_l of [1] have the values 6 and -18 for K^2.

References

1. S. S. Chern, *On Riemannian manifolds of four dimensions*, Bull. Amer. Math. Soc. vol. 51 (1943).
2. ———, *Characteristic classes of Hermitian manifolds*, Ann. of Math. vol. 47 (1946).
3. C. Chevalley, *Theory of Lie groups* I, Princeton, 1946.
4. H. Hopf and M. Rueff, *Über faserungstreue Abbildungen der Sphären*, Comment. Math. Helv. vol. 11 (1938–1939).
5. N. E. Steenrod, *The classification of sphere bundles*, Ann. of Math. vol. 45 (1944).
6. R. Courant, *Differential- and integral calculus* II, New York, 1936.

University of Michigan

FINITELY-VALUED COHOMOLOGY GROUPS

JOHN W. KEESEE[1]

In this note it will be shown that finitely-valued functions are adequate to determine the Alexander-Kolmogoroff cohomology groups of a compact Hausdorff space. This is effected by exhibiting a homomorphism of the finitely-valued groups into the general groups which is an isomorphism onto if the space is compact. We assume familiarity with Spanier's [2][2] development of the Alexander-Kolmogoroff groups although the definitions are repeated here.

Let X be an arbitrary topological space and G a fixed abelian group. Denote by $\Phi^p(X)$ the group of all functions (p-functions) from X^{p+1} (the Cartesian product of X with itself $p+1$ times) to G. Let $\{N(\Delta X^{p+1})\}$ represent the open sets containing the diagonal of X^{p+1}. Then $\Phi_0^p(X)$ is the subgroup of $\Phi^p(X)$ consisting of those p-functions which are zero on some $N(\Delta X^{p+1})$. There is a homomorphism $\bar{\delta}$ of $\Phi^p(X)$ into $\Phi^{p+1}(X)$ defined by the formula

$$\bar{\delta}\phi(x_0, x_1, \cdots, x_{p+1}) = \sum_{i=0}^{p+1} (-1)^i \phi(x_0, x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{p+1}).$$

This homomorphism has the property that $\bar{\delta}\bar{\delta} = 0$. The subgroup $\Phi_Z^p(X) = \bar{\delta}^{-1}\Phi_0^{p+1}(X)$ is called the group of p-cocycles, and $\Phi_B^p(X)$, the group of p-coboundaries, is defined to be $\bar{\delta}\Phi^{p-1}(X) + \Phi_0^p(X)$. The pth cohomology group of the space X, denoted by $H^p(X)$, is the quotient group $\Phi_Z^p(X)/\Phi_B^p(X)$.

For each $f: X \to Y$ (f not necessarily continuous) there is defined a homomorphism $f^\sharp: \Phi^p(Y) \to \Phi^p(X)$ by

$$f^\sharp\phi(x_0, x_1, \cdots, x_p) = \phi(f(x_0), f(x_1), \cdots, f(x_p)), \quad \phi \in \Phi^p(Y).$$

If f is a map (=continuous function), $f^\sharp$ carries each of the subgroups $\Phi_0^p(Y)$, $\Phi_B^p(Y)$, and $\Phi_Z^p(Y)$ into the corresponding subgroup of $\Phi^p(X)$.

Let A be a subset of X and $i: A \to X$ the identity map. Then

$$\Phi^p(X, A) = i^{\sharp-1}\Phi_0^p(A), \qquad \Phi_0^p(X, A) = \Phi_0^p(X),$$
$$\Phi_B^p(X, A) = \bar{\delta}\Phi^{p-1}(X, A) + \Phi_0^p(X, A),$$
$$\Phi_Z^p(X, A) = \Phi^p(X, A) \cap \Phi_Z^p(X).$$

Presented to the Society, April 2, 1949; received by the editors June 3, 1949.

[1] This work was done under Contract N7-onr-534, Task Order III, Navy Department, Office of Naval Research.

[2] Numbers in brackets refer to the references cited at the end of the paper.

The pth cohomology group of X mod A is $H^p(X, A) = \Phi_Z^p(X, A)/\Phi_B^p(X, A)$.

If f is a function on X to Y that carries a subset A of X into a subset B of Y, f is said to be on (X, A) to (Y, B). We write

$$f: \quad (X, A) \to (Y, B).$$

If $f:(X, A)\to(Y, B)$ is a map, $f^\sharp$ carries $\Phi^p(Y, B)$, $\Phi_Z^p(Y, B)$, and $\Phi_B^p(Y, B)$ into $\Phi^p(X, A)$, $\Phi_Z^p(X, A)$, and $\Phi_B^p(X, A)$ respectively.

For a fixed pair of functions f and g on (X, A) to (Y, B) there is a function D: $\Phi^{p+1}(Y, B)\to\Phi^p(X, A)$ defined by $D\phi(x_0, x_1, \cdots, x_p) = \sum_{i=0}^{p} (-1)^i\phi(g(x_0), g(x_1), \cdots, g(x_i), f(x_i), \cdots, f(x_p))$. Then for $p>0$ and $\phi\in\Phi^p(Y, B)$, $f^\sharp\phi - g^\sharp\phi = \bar{\delta}D\phi + D\bar{\delta}\phi$. If $p=0$, $f^\sharp\phi - g^\sharp\phi = D\bar{\delta}\phi$.

The purpose of this paper is to consider the groups $H_F^p(X, A)$ defined as follows. Let $F^p(X)$ be the subgroup of $\Phi^p(X)$ consisting of those p-functions which have only a finite number of elements of G as values. Define $F^p(X, A) = \Phi^p(X, A)\cap F^p(X)$, $F_0^p(X, A) = \Phi_0^p(X, A) \cap F^p(X)$, $F_B^p(X, A) = \bar{\delta}F^{p-1}(X, A) + F_0^p(X, A)$ (when $p=0$, $F_B^p(X, A) = 0$), $F_Z^p(X, A) = \Phi_Z^p(X, A)\cap F^p(X)$. Then $H_F^p(X, A)$ is the quotient group $F_Z^p(X, A)/F_B^p(X, A)$. It will be shown that for an arbitrary space X, there is a homomorphism $\tau: H_F^p(X, A)\to H^p(X, A)$ which is an isomorphism onto if X is a compact Hausdorff space and A is closed.

LEMMA 1. *Let M and N be abelian groups with subgroups M_0 and N_0 and λ a homomorphism of M into N such that $\lambda(M_0)\subset N_0$. Let $\alpha: M\to M/M_0$ and $\beta: N\to N/N_0$ be the natural homomorphisms. Then there is a homomorphism $\rho: M/M_0\to N/N_0$ defined by $\rho\alpha(x) = \beta\lambda(x)$ for $x\in M$. Furthermore, the kernel of ρ is $\lambda^{-1}(N_0)/M_0$ and the image of M/M_0 under ρ is $(\lambda(M)+N_0)/N_0$.*

The method of proof is evident.

The results of Lemma 1 are used to define the homomorphism τ as follows: Denote by $\gamma: \Phi_Z^p(X, A)\to H^p(X, A)$ and $\nu: F_Z^p(X, A) \to H_F^p(X, A)$ the natural homomorphisms and let $\sigma: F_Z^p(X, A) \to \Phi_Z^p(X, A)$ be the identity homomorphism. Define $\tau: H_F^p(X, A) \to H^p(X, A)$ by $\tau\nu(\phi) = \gamma\sigma(\phi)$ for $\phi\in F_Z^p(X, A)$. It follows that τ is a homomorphism, the kernel of τ is $\Phi_B^p(X, A)\cap F^p(X)/F_B^p(X, A)$, and the image of $H_F^p(X, A)$ under τ is $F_Z^p(X, A)+\Phi_B^p(X, A)/\Phi_B^p(X, A)$.

The next two lemmas are concerned with properties of subsets of G that satisfy the following two conditions: (1) If a subset H of G has property M, then $(-H)$ has M. (2) If H_1 and H_2 have M, then (H_1+H_2) has M. For such a property M, define $M^p(X, A)$ to be all

$\phi \in \Phi^p(X, A)$ such that ϕ carries some $N(\Delta X^{p+1})$ into a set H with property M.

In what follows, if $\mathcal{U}$ is a covering of X, $N_p(\mathcal{U})$ will denote the set $\cup[U^p \mid U \in \mathcal{U}]$.

LEMMA 2. *Hypothesis*: (a) *X and Y are topological spaces, f and g are functions on (X, A) to (Y, B);*

(b) $\phi \in \Phi^p(Y, B), \bar{\delta}\phi \in M^{p+1}(Y, B)$;

(c) *$\mathcal{V}$ is a covering of Y such that $\phi = 0$ on $N_{p+1}(\mathcal{V}) \cap B^{p+1}$, and $\bar{\delta}\phi$ carries $N_{p+2}(\mathcal{V})$ into a subset N of G with property* M;

(d) *$\mathcal{U}$ is a covering of X such that for each $U \in \mathcal{U}$, there is a $V \in \mathcal{V}$ for which $f(U) \cup g(U) \subset V$.*

Conclusion: *For $p > 0$, $f^\sharp\phi - g^\sharp\phi \in \bar{\delta}\Phi^{p-1}(X, A) + M^p(X, A)$; for $p = 0$, $f^\sharp\phi - g^\sharp\phi \in M^p(X, A)$.*

PROOF. It is sufficient to show (i) if $p \geqq 0$, $D\bar{\delta}\phi \in M^p(X, A)$ and (ii) if $p > 0$, $\bar{\delta}D\phi \in \bar{\delta}\Phi^{p-1}(X, A)$. First consider $(x_0, \cdots, x_p) \in N_{p+1}(\mathcal{U}) \cap A^{p+1}$. Then $D\bar{\delta}\phi(x_0, \cdots, x_p) = \sum_{i=0}^{p} (-1)^i(\bar{\delta}\phi)(g(x_0), \cdots, g(x_i), f(x_i), \cdots, f(x_p))$. But $(g(x_0), \cdots, g(x_i), f(x_i), \cdots, f(x_p)) \in N_{p+2}(\mathcal{V}) \cap B^{p+2}$ so $D\bar{\delta}\phi = 0$ on $N_{p+1}(\mathcal{U}) \cap A^{p+1}$ and therefore $D\bar{\delta}\phi \in \Phi^p(X, A)$. Also for $(x_0, \cdots, x_p) \in N_{p+1}(\mathcal{U})$, $D\bar{\delta}\phi(x_0, \cdots, x_p) \in \sum_{i=0}^{p}(-1)^i N$. But this set has property M since N has and $D\bar{\delta}\phi \in M^p(X, A)$. Second, for $p > 0$, take $(x_0, \cdots, x_{p-1}) \in N_p(\mathcal{U}) \cap A^p$. Then $(g(x_0), \cdots, g(x_i), f(x_i), \cdots, f(x_{p-1})) \in N_{p+1}(\mathcal{V}) \cap B^{p+1}$ and $D\phi(x_0, \cdots, x_{p-1}) = 0$. So $D\phi \in \Phi^{p-1}(X, A)$. This completes the proof of the lemma.

LEMMA 3. *If X is a compact Hausdorff space and A is closed, then* $\Phi^p(X, A) \cap \bar{\delta}^{-1}M^{p+1}(X, A) \subset \bar{\delta}\Phi^{p-1}(X, A) + M^p(X, A) + F^p(X, A)$.

PROOF. If $\phi \in \Phi^p(X, A) \cap \bar{\delta}^{-1}M^{p+1}(X, A)$, then there exists a covering $\mathcal{V}$ of X for which $\phi = 0$ on $N_{p+1}(\mathcal{V}) \cap A^{p+1}$ and $\bar{\delta}\phi$ carries $N_{p+2}(\mathcal{V})$ into a subset of G with property M. Let $\mathcal{U}$ be a finite star refinement of $\mathcal{V}$ [1, p. 324]. Write $\mathcal{U} = [U_1, U_2, \cdots, U_n]$. To each U_j for which $U_j \cap A \neq \square$, assign $\bar{x}_j \in U_j \cap A$. To each U_j such that $U_j \cap A = \square$, assign $\bar{x}_j \in U_j$. Define $f: (X, A) \to (X, A)$ by $f(x) = \bar{x}_j$ where j is the least integer for which $x \in U_j$. Then $f(X)$ is a finite subset of X and $f^\sharp\phi \in F^p(X)$. Also $f^\sharp\phi \in F^p(X, A)$, for consider $(x_0, \cdots, x_p) \in N_{p+1}(\mathcal{U}) \cap A^{p+1}$. Then $x_i \in U_k$ $(i = 0, \cdots, p)$ and there exists $V_0 \in \mathcal{V}$ such that $V_0 \supset$ the star of U_k. Then $f(U_k) \subset$ the star of U_k implies $f(x_i) \in$ the star of U_k. Hence $(f(x_0), f(x_1), \cdots, f(x_p)) \in V_0^{p+1}$, and $f^\sharp\phi(x_0, \cdots, x_p) = \phi(f(x_0), \cdots, f(x_p)) = 0$. Therefore $f^\sharp\phi \in \Phi^p(X, A)$. Let $i: (X, A) \to (X, A)$ be the identity function; then for each $U \in \mathcal{U}$, there exists $V \in \mathcal{V}$ such that $i(U) \cup f(U) \subset V$. Now by Lemma 2, $i^\sharp\phi - f^\sharp\phi \in \bar{\delta}\Phi^{p-1}(X, A) + M^p(X, A)$ or

$$\phi \in \bar{\delta}\Phi^{p-1}(X, A) + M^p(X, A) + F^p(X, A).$$

For the next lemma define $L^p(X)$ to be the group of p-functions which assume only a finite number of values on some $N(\Delta X^{p+1})$, and let $L^p(X, A) = L^p(X) \cap \Phi^p(X, A)$. It follows that $L^p(X, A) = F^p(X, A) + \Phi_0^p(X, A)$ and $\bar{\delta}L^p(X, A) \subset \bar{\delta}F^p(X, A) + \Phi_0^{p+1}(X, A)$.

LEMMA 4. *If X is a compact Hausdorff space and A is closed, then* $\Phi^p(X, A) \cap \bar{\delta}^{-1}L^{p+1}(X, A) \subset \Phi_B^p(X, A) + F^p(X, A)$.

PROOF. Since the property of being finite is a property of the type of Lemma 2, it follows by Lemma 3 that

$$\Phi^p(X, A) \cap \bar{\delta}^{-1}L^{p+1}(X, A) \subset \bar{\delta}\Phi^{p-1}(X, A) + L^p(X, A).$$

But $\bar{\delta}\Phi^{p-1}(X, A) + L^p(X, A) = \bar{\delta}\Phi^{p-1}(X, A) + \Phi_0^p(X, A) + F^p(X, A) = \Phi_B^p(X, A) + F^p(X, A)$.

In the proof of the following theorem, repeated use is made of the so-called Dedekind property of groups, that is, if A, B, and C are subgroups of a group G, and A is contained in C, then

$$(A + B) \cap C = (A \cap C) + (B \cap C).$$

THEOREM. *If X is a compact Hausdorff space and A is closed, then τ is an isomorphism onto.*

PROOF. (a) Show τ is an isomorphism. The kernel of τ is $\Phi_B^p(X, A) \cap F^p(X)/F_B^p(X, A)$.

Now

$$\begin{aligned}\Phi_B^p(X, A) \cap F^p(X) &= [\bar{\delta}\Phi^{p-1}(X, A) + \Phi_0^p(X, A)] \cap F^p(X) \\ &= [\bar{\delta}\Phi^{p-1}(X, A) + \Phi_0^p(X, A)] \cap L^p(X) \cap F^p(X) \\ &= [\bar{\delta}\Phi^{p-1}(X, A) \cap L^p(X, A) + \Phi_0^p(X, A)] \cap F^p(X).\end{aligned}$$

By Lemma 4 $\bar{\delta}\Phi^{p-1}(X, A) \cap L^p(X, A) \subset \bar{\delta}\Phi_B^{p-1}(X, A) + \bar{\delta}F^{p-1}(X, A)$. Therefore, $\Phi_B^p(X, A) \cap F^p(X) \subset [\bar{\delta}F^{p-1}(X, A) + \Phi_0^p(X, A)] \cap F^p(X)$. But $[\bar{\delta}F^{p-1}(X, A) + \Phi_0^p(X, A)] \cap F^p(X) = \bar{\delta}F^{p-1}(X, A) + \Phi_0^p(X, A) \cap F^p(X) = F_B^p(X, A)$. This shows that the kernel of τ is 0.

(b) Show τ is onto. The image of $H_F^p(X, A)$ under τ is $F_Z^p(X, A) + \Phi_B^p(X, A)/\Phi_B^p(X, A)$. In Lemma 3 define "$H$ has M" to mean "$H = 0$." Then $M^{p+1}(X, A) = \Phi_0^{p+1}(X, A)$ and

$$\begin{aligned}\Phi_Z^p(X, A) &\subset \bar{\delta}\Phi^{p-1}(X, A) + \Phi_0^p(X, A) + F^p(X, A) \\ &= \Phi_B^p(X, A) + F^p(X, A).\end{aligned}$$

But since $\Phi_B^p(X, A) \subset \Phi_Z^p(X, A)$, it follows that $\Phi_Z^p(X, A) \subset \Phi_B^p(X, A)$

$+F_2^q(X, A)$. The reverse inclusion is clear and the equality shows that τ is onto.

References

1. S. Lefschetz, *Algebraic topology*, Amer. Math. Soc. Colloquium Publications, vol. 27, New York, 1942.

2. E. H. Spanier, *Cohomology theory for general spaces*, Ann. of Math. vol. 49 (1948) pp. 405–427.

Tulane University

A NOTE ON A COMBINATORIAL PROBLEM

H. J. RYSER

The purpose of this note is to prove the following theorem.

THEOREM. *Let $x_1, x_2, \cdots, x_v$ denote v distinct elements and let $T_1, T_2, \cdots, T_v$ denote v sets formed from these elements. Suppose further that each T_i consists of exactly k distinct elements and that every pair of distinct sets T_i and T_j have exactly λ elements in common $(0<\lambda<k<v)$. Then*

$$\lambda = \frac{k(k-1)}{v-1}.$$

The combinatorial problem under consideration makes its appearance in the study of projective planes, Hadamard matrices, and block designs. In the combinatorial problem of Todd arising in the study of Hadamard matrices it is usually assumed that $\lambda = k(k-1)/(v-1)$ [3; 4].[1] For the symmetrical block designs each element of the arrangement is required to occur exactly k times, and it is then easy to verify that $\lambda = k(k-1)/(v-1)$ [1]. Further theorems concerning the possibility of this combinatorial problem for a given v, k, and λ may be found in [3].

To prove the theorem let the elements $x_1, \cdots, x_v$ be listed in a row, and let the sets $T_1, \cdots, T_v$ be listed in a column. Form the incidence matrix A of the arrangement in the usual way by inserting a one in column i and row j if x_i belongs to set T_j, and a zero in the contrary case [2; 3]. Now let s_i denote the sum of column i of the

Presented to the Society, September 2, 1949; received by the editors June 9, 1949.

[1] Numbers in brackets refer to the references cited at the end of the paper.

matrix A. Let row r of A have ones in the k columns $r_1<r_2<\cdots<r_k$, and let $t=\lambda(v-1)+k$. Then by the hypothesis of the theorem it follows that

$$s_{r_1}+s_{r_2}+\cdots+s_{r_k}-t=0.$$

The matrix of the above system of v equations in v unknowns is simply A. Moreover, it is clear that

$$s_1+s_2+\cdots+s_v-kv=0.$$

Now $t\neq 0$, and consequently the matrix

$$C=\begin{bmatrix} & & 1 \\ & A & \vdots \\ & & 1 \\ 1\cdots 1 & & kv/t \end{bmatrix}$$

has a zero determinant. Multiplying C by its transpose implies that

$$\begin{vmatrix} k+1 & \lambda+1 & \cdots & \lambda+1 & r \\ \lambda+1 & k+1 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \lambda+1 & & & k+1 & r \\ r & & \cdots & r & s \end{vmatrix}=0,$$

where $r=k+kv/s$ and $s=v+k^2v^2/t^2$. Multiplying row $v+1$ by $-r/s$ and adding to each of the other rows gives $s\cdot\det D=0$, where the matrix D has $k+1-r^2/s$ in the main diagonal and $\lambda+1-r^2/s$ in all other positions. An evaluation of the determinant of the matrix D implies that

$$(k-\lambda)^{v-1}\{k+1-r^2/s+(v-1)(\lambda+1-r^2/s)\}=0,$$

and consequently

$$t+v-r^2v/s=0.$$

Since

$$\frac{r^2v}{s}=\frac{k^2(t+v)^2}{t^2+k^2v},$$

it is clear that

$$(t+v)(t^2+k^2v)=k^2(t+v)^2.$$

Thus $t^2+k^2v=k^2(t+v)$ and $t=k^2$.

COROLLARY. *Every element in the arrangement described in the preceding theorem occurs exactly k times and every pair of elements occurs exactly* λ *times.*

The corollary is an immediate consequence of the preceding theorem and Theorem 1 derived in [3].

A simple and direct alternative proof of the corollary may be obtained by noting that for $\lambda=k(k-1)/(v-1)$, the equations $s_1+\cdots+s_v=kv$ and $s_1(s_1-1)+\cdots+s_v(s_v-1)=\lambda v(v-1)$ imply that $s_i=k$ (see [1]). The writer is indebted to Professor Marshall Hall, Jr., for this suggestion.

BIBLIOGRAPHY

1. R. C. Bose, *On the construction of balanced incomplete block designs*, Annals of Eugenics vol. 9 (1939) pp. 353–399.

2. R. H. Bruck and H. J. Ryser, *The nonexistence of certain finite projective planes*, Canadian Journal of Mathematics vol. 1 (1949) pp. 88–93.

3. S. Chowla and H. J. Ryser, *Combinatorial problems*, Canadian Journal of Mathematics vol. 2 (1950) pp. 93–99.

4. J. A. Todd, *A combinatorial problem*, Journal of Mathematics and Physics vol. 12 (1933) pp. 321–333.

THE INSTITUTE FOR ADVANCED STUDY

LINEAR REPRESENTATIONS OF TOPOLOGICAL GROUPS

MORIKUNI GOTÔ

Introduction. In his recent papers entitled *Faithful representations of Lie groups*, I and II,[1] the author has called a locally compact connected group L *faithfully representable* (*f.r.*) if L admits a faithful representation,[2] and has determined the characteristic properties of f.r. groups. It is to be pointed out that although an f.r. group is a Lie group, a connected Lie group is not always f.r.[3] Furthermore, he introduced there a notion of (l)-groups, and proved that *a Lie group is f.r. if it is an* (l)*-group*,[4] *where the definition of* (l)*-groups runs as* follows:

DEFINITION 1. *Let G be a locally compact connected group. Then G is called an (l)-group if for every element g of G distinct from the identity there exists a representation of G under which g does not go over into the unit matrix.*

The above definition is obviously equivalent to the one of imposing on G the condition that G *can be approximated by f.r. groups*, namely that there exists a system $\{N_\alpha\}$ of closed invariant subgroups of G so that every factor group G/N_α is f.r. and the intersection of all N_α's coincides with the identity group $\{e\}$, which is composed of the identity e only.

In the present paper we shall study the structure of (l)-groups as a sequel of [M.G.]. Now K. Iwasawa called a locally compact group G an (L)*-group* if G *can be approximated by Lie groups*; he studied the structure of (L)-groups, and in particular solved the so-called fifth problem of Hilbert for such groups.[5] Our (l)-group is of course a connected (L)-group and our study is based naturally on the theory of (L)-groups also.

Next we shall state the outline of the present paper. In §1 we study *semi-simple* (L)*-groups*, where the semi-simpleness of an (L)-group is defined in an analogous way as in the theory of Lie groups. §2 is

Received by the editors May 3, 1949.

[1] Mathematica Japonicae vol. 1 (1949), and Nagoya Mathematical Journal vol. 1 (1950). Referred to as [M.G.], or explicitly [M.G.I.] or [M.G.II.].

[2] By a *representation* we mean a continuous linear representation. An f.r. group can also be defined as a topological group topologically isomorphic with a *closed* connected subgroup of the general linear group. [M.G.II, Theorem 9].

[3] See, for example [M.G.].

[4] See [M.G.II, Theorem 8].

[5] K. Iwasawa, *On some types of topological groups*, Ann. of Math. vol. 50 (1949). We shall refer to this paper as [K.I.].

devoted to the proof of Theorem 2, which has an essential bearing on the latter part of this paper. Now in §3 we deal with the class of *completely reducible* (l)*-groups*, which contains both compact connected groups and semi-simple (l)-groups. After such preparations we shall investigate in §4 the structure of general (l)-groups. Theorem 5, as well as Theorem 6, is our fundamental result, affording a characterization of (l)-groups. Lastly in §5 some supplementary facts about solvable (l)-groups are stated.

In closing the introduction, the author wishes to express his sincere gratitude to Mr. Y. Matsushima for his valuable suggestions and cooperation.

1. **Semi-simple (L)-groups.** Let G be a locally compact group. G contains then the *radical*, that is, the uniquely determined maximal connected solvable invariant subgroup, which is closed in G in virtue of [K.I.]. Now we shall extend the notion of semi-simpleness of Lie groups to (L)-groups, as follows:

DEFINITION 2. *An (L)-group is called semi-simple if its radical coincides with the identity group.*

Since the radical is characteristic, any closed invariant subgroup of a semi-simple (L)-group is also semi-simple.

Now in order to study the structure of semi-simple (L)-groups we first give the following two lemmas.

LEMMA 1. *Let G be a locally compact connected group, and let C and N be closed invariant subgroups of G such that $G=CN$. Denote by C^0 and N^0 the connected components of C and N containing the identity respectively. Suppose C, or C^0, be compact, then we have $G=C^0N^0$.*

PROOF. As C^0 is compact, C^0N is a closed invariant subgroup. Since the intersection of C^0N and C contains C^0, the isomorphism[6] $G/C^0N \cong C/C^0N\cap C$ implies that G/C^0N is 0-dimensional. On the other hand, G/C^0N must be connected because of the connectedness of G. Hence we get $G=C^0N$. Now a similar argument as above would establish the lemma. Q.E.D.

LEMMA 2. *Let G be a connected semi-simple (L)-group, and K a compact 0-dimensional invariant subgroup of G. Suppose that the factor group G/K is a Lie group. Then G itself is a Lie group.*

PROOF. It is clearly sufficient to prove that K cannot be infinite and separable. Suppose that K be an infinite separable group. Then

[6] By an *isomorphism* (or a *homomorphism*) we mean a continuous open isomorphism (or homomorphism).

there exists a sequence $\{K_n\}$:

$$K = K_1 \supset K_2 \supset K_3 \supset \cdots, \qquad K_n \neq K_{n+1},$$

of open subgroups of K such that $\bigcap K_n = \{e\}$. Since K is a central invariant subgroup and any open subgroup of a topological group is closed, K_n's are all closed invariant subgroups.

We denote by $\bar{G}$ the universal covering group of G/K. Then we can readily select a sequence $\{Z_n\}$ of central invariant subgroups of $\bar{G}$:

$$Z_1 \supset Z_2 \supset Z_3 \supset \cdots, \qquad Z_n \neq Z_{n+1},$$

such that $\bar{G}/Z_n \cong G/K_n$, $n = 1, 2, \cdots$. Denote by Z the intersection of all Z_n. Then it is easy to construct an algebraic isomorphism between $\bar{G}/Z$ and G. Since K is central, K is a countable group. On the other hand, it is easy to prove that a countable compact group is discrete. Hence K must be finite, contrary to the hypothesis. This completes our proof.

Now let G be a connected semi-simple (L)-group and let C^* be the (uniquely determined) maximal compact invariant subgroup[7] of G. Then by a theorem of Iwasawa in [K.I.], G contains a closed invariant subgroup L^* of G such that $G = C^*L^*$ and $[C^*, L^*] = \{e\}$, where we denote by $[C^*, L^*]$ the subgroup generated by the elements of the form $c^{-1}l^{-1}cl$ with $c \in C^*$ and $l \in L^*$. Let C and L be the connected components of C^* and L^* respectively containing the identity. Then, using Lemma 1, we obtain that $G = CL$.

Let now K be the maximal compact invariant subgroup of L, and K^0 the connected component of K containing e. C being the maximal connected compact invariant subgroup, K^0 must be contained in the intersection of C and L. Hence K^0 is central. This shows that K^0 coincides with the identity group, whence K is 0-dimensional. Therefore, using Lemma 2, we get the fact that L is a Lie group. Now denote by D the intersection of C and L. Since D is compact and is contained in the center of L, D must be a finite group.

Next let $\{\bar{C}_\lambda\}$ be a sequence of compact connected simply connected semi-simple simple Lie groups, and $\bar{C}$ the direct product of all $\bar{C}_\lambda$'s: $\bar{C} = \prod \bar{C}_\lambda$. $\bar{C}$ is also a compact group with respect to the so-called weak topology. Now let D be a closed 0-dimensional invariant subgroup of $\bar{C}$. Then the factor group $\bar{C}/D$ is obviously semi-simple, and conversely any connected compact semi-simple group is readily

[7] The existence, as well as the uniqueness, of such a subgroup for any connected (L)-group has been assured in [K.I.].

proved to be obtained in such a way.[8]

In a similar way, we can establish a method of constructing all connected semi-simple (L)-groups: Let $\{\bar{S}_\lambda\}$ be a sequence of connected simply connected semi-simple simple Lie groups such that almost all of $\bar{S}_\lambda$'s are compact, and let $\bar{S}$ be the direct product of all $\bar{S}_\lambda$'s: $\bar{S}=\prod\bar{S}_\lambda$. Then $\bar{S}$ is clearly a connected semi-simple (L)-group with respect to the weak topology. Now let D be a closed 0-dimensional invariant subgroup of $\bar{S}$. $\bar{S}/D$ is clearly semi-simple, and *any connected semi-simple (L)-group is isomorphic with some $\prod\bar{S}_\lambda/D$ for suitable $\{\bar{S}_\lambda\}$ and D.*

Now a factor group of the above $\prod\bar{S}_\lambda$ is clearly semi-simple. Hence any factor group of a connected semi-simple (L)-group is also semi-simple. Thus we get the following theorem.

THEOREM 1. (1) *Let G be a connected semi-simple (L)-group, and C the maximal compact connected invariant subgroup of G. Then there exists a closed connected invariant Lie subgroup*[9] *L of G so that $G=CL$, and $[C, L]=\{e\}$, and that the intersection of C and L is a finite group.*

(2) *Any factor group of G is semi-simple.*

Using the latter part of Theorem 1, we can easily prove the following corollary:

COROLLARY 1. *Let G be a connected (L)-group and R the radical of G. If N is a closed invariant subgroup of G such that RN is closed, then RN/N is the radical of G/N.*

COROLLARY 2. *Let G be a connected group and S a closed connected invariant subgroup. We denote by $S^\perp$ the centralizer*[10] *of S. If S is a semi-simple (L)-group, then we have $G=SS^\perp$.*

PROOF. Corollary 2 has been proved in [K.I.] for any compact invariant subgroup S. It is also clear for a connected semi-simple Lie group S. Now it is not hard to establish the corollary from these special cases. Q.E.D.

2. **Proof of Theorem 2.** In this section we shall give a proof of the following theorem.

THEOREM 2. *Let G be an (l)-group and K a compact invariant subgroup. Then the factor group G/K is also an (l)-group.*

[8] Cf. H. Freudenthal, *Topologische Gruppen mit genügend vielen fastperiodischen Funktionen*, Ann. of Math. vol. 37 (1936).

[9] By a (closed) *Lie subgroup* we mean a subgroup which is a Lie group.

[10] Namely, the (invariant) subgroup composed of all elements which commute with every element of S.

In order to prove Theorem 2, we first recollect and modify some results stated in [M.G.].

Let G be an f.r. group, and N a maximal closed simply connected solvable invariant subgroup of G. Then there exists a closed subgroup T, whose radical is compact, such that

$$G = TN, \qquad T \cap N = \{e\}.$$

Conversely a connected Lie group is f.r. if it admits such a decomposition.[11]

Now the radical of an f.r. group is compact if and only if it is completely reducible,[12] where the complete reducibility of an (l)-group is defined as follows:

DEFINITION 3. *An (l)-group G is called completely reducible if every representation of G is completely reducible.*

Let T be a completely reducible f.r. group and A the radical of T. Then A is compact and there exists in T a closed connected semi-simple subgroup S such that $T=SA$. It is clear that $[S, A]=\{e\}$ and the intersection of S and A is a finite group. Now any factor group of a semi-simple f.r. group is also f.r.,[13] and a connected Lie group is f.r. if its radical and a maximal semi-simple subgroup[14] are f.r.[15] These obviously imply that a factor group of a completely reducible f.r. group is also a completely reducible f.r. group.

Next, it is to be remarked that a simply connected solvable Lie group is homeomorphic with the Euclidean space, and that a locally compact connected solvable group is a simply connected Lie group if and only if it contains no compact subgroup other than the identity group.[16]

Now the following lemma is a special case of Theorem 2.

LEMMA 3. *Let G be an f.r. group and K a compact invariant subgroup. Then the factor group G/K is also f.r.*

PROOF. Let $G=TN$ be a decomposition as above, and let M be a maximal compact subgroup[17] of T. Since N is homeomorphic with

[11] [M.G.II, Corollary 2 to Theorem 7].

[12] [M.G.II, Lemma 14].

[13] [M.G.II, Lemma 9].

[14] By a *maximal semi-simple subgroup* of a Lie group we mean a connected group generated by a closed maximal semi-simple local subgroup.

[15] [M.G.II, Theorem 7]. A. Malcev, *On linear Lie groups*, C.R. (Doklady) Acad. Sci. URSS. (1943).

[16] See [K.I.].

[17] See [K.I.].

the Euclidean space, M is maximal compact also in G. Hence K is contained in M, and thus we have $K \subset T$. Now since T is completely reducible, the factor group T/K is a completely reducible f.r. group. It is now clear that G/K is f.r., because

$$G/K = T/K \cdot (NK)/K$$

is also a decomposition as above. Q.E.D.

PROOF OF THEOREM 2. Let g be an element of G which is not contained in K. For any element x of K we can readily select a representation f_x of G and a neighborhood U_x of x such that $f_x(U_x g)$ does not contain the unit matrix. Now there exists a finite system $\{x_1, \cdots, x_n\}$ of elements of K such that the system $\{U_{x_1}, \cdots, U_{x_n}\}$ of neighborhoods covers K: $\bigcup U_{x_i} \supset K$.

Now let f be the so-called sum representation of $f_{x_1}, \cdots, f_{x_n}$, and let $N_1, N_2, \cdots, N_n, N$ be the kernel of $f_{x_1}, f_{x_2}, \cdots, f_{x_n}, f$. From the relations $U_{x_i} g \cap N_i = 0$ $(i = 1, 2, \cdots, n)$, and $\bigcap N_i = N$, we have $U_{x_i} g \cap N = 0$, where 0 denotes the empty set. Accordingly we obtain $Kg \cap N = 0$, whence

$$Kg \cap KN = 0.$$

Now since G/N is f.r. and KN/N is a compact invariant subgroup of G/N, G/KN is also f.r. by Lemma 3. Now the relation $Kg \cap KN = 0$ implies that Kg/K is not contained in the kernel of the homomorphism $G/K \sim G/KN : G/K/KN/K \cong G/KN$, and this completes our proof.

COROLLARY. *Let G be an (l)-group, and U a neighborhood of the identity in G. Then there exists a compact invariant subgroup K, which is contained in U, such that the factor group G/K is f.r.*

PROOF. Since G is a connected (L)-group, there exists a compact invariant subgroup K such that G/K is a Lie group and $K \subset U$, in virtue of [K.I.]. Thus the corollary is an immediate consequence of Theorem 2. Q.E.D.

3. **Completely reducible (l)-groups.** In [M.G. II][18] we have determined the class of all semi-simple f.r. groups. Here we consider semi-simple (l)-groups.

THEOREM 3. (1) *Let G be a semi-simple (L)-group, and let*

$$G = CL$$

[18] Lemma 9, Theorem 3, and Theorem 4.

be the decomposition of G as in Theorem 1. *Then G is an (l)-group if and only if L is f.r.*

(2) *The center of a semi-simple (l)-group is compact.*

(3) *Any factor group of a semi-simple (l)-group is also an (l)-group.*

PROOF. Assume that L be f.r. Then the direct product $C \times L$ is clearly an (l)-group. Now we can readily find a finite central invariant subgroup D of $C \times L$ such that $C \times L/D \cong G$. Hence G is an (l)-group by Theorem 2. The converse is trivial.

Next, we know that the center of a semi-simple f.r. group is finite.[19] From this the second assertion follows immediately.

Further, let G be a semi-simple (l)-group and N a closed invariant subgroup of G. We denote by N^0 the connected component of N containing the identity. N^0 is clearly a semi-simple (l)-group. Let C^0 be the maximal connected compact invariant subgroup of N^0 and let

$$N^0 = C^0 L^0$$

be the decomposition of N^0 as in Theorem 1. C^0 is obviously an invariant subgroup of G, and the compactness of C^0 implies that G/C^0 is an (l)-group by Theorem 2. Let now $G/C^0 = C^* L^*$ be the decomposition of G/C^0 as in Theorem 1. Then the invariant subgroup N^0/C^0 is clearly contained in L^*. The fact that a factor group of a semi-simple f.r. group is f.r. readily implies that G/N^0 is also an (l)-group. Hence the center of G/N^0 is compact by our above assertion. This shows the compactness of N/N^0, and we obtain the last assertion, using Theorem 2 again. Q.E.D.

It is well known that a semi-simple f.r. group is completely reducible, and so is a semi-simple (l)-group. Now, we study general completely reducible (l)-groups. A criterion for the complete reducibility is furnished by the following lemma.

LEMMA 4. *An (l)-group is completely reducible if and only if its radical is compact.*

PROOF. Let G be a completely reducible (l)-group. According to the corollary of Theorem 2, there exists a compact invariant subgroup K such that G/K is f.r. Denote by R the radical of G. Then RK/K is the radical of G/K, in virtue of Corollary 1 to Theorem 1. Now since G/K is obviously completely reducible, its radical RK/K is compact. Hence R must be compact.

Conversely, let G be an (l)-group with the compact radical R, and let N be an arbitrary closed invariant subgroup of G such that the

[19] [M.G.I, Lemma 5].

factor group G/N is f.r. It is sufficient to show that G/N is completely reducible, or what is the same, that the radical of G/N is compact. Now the compactness of R implies that RN/N is the compact radical, and this completes our proof.

Let G be a topological group. We denote by $d(G)$ the commutator group of G, that is, $[G, G]$, and by $D(G)$ the closure of $d(G)$. Following Iwasawa we call $D(G)$ the *topological commutator group* of G. The topological commutator group of a connected group is obviously connected.

Next, let G be a completely reducible (l)-group and R the radical of G. Since R is compact, $D(G)R/D(G)$ is the radical of the commutative $G/D(G)$. Hence $D(G)R$ coincides with G: $G=D(G)R$. Now let K be a compact invariant subgroup of G. Since it is clear that $\overline{d(G)K} = D(G)K$, we have

$$D(G)K/K = D(G/K).$$

Let Z be the intersection of $D(G)$ and R. In any representation of G with a compact kernel, Z goes over into the intersection of the commutator group and the radical, which is obviously a finite group. Hence Z must be 0-dimensional. That $D(G)$ is semi-simple follows immediately.

Now it is easy to prove the following theorem.

THEOREM 4. (1) *Let G be a completely reducible (l)-group, and C the maximal connected compact invariant subgroup. G contains a closed connected semi-simple f.r. invariant subgroup L such that $G=CL$, $[C, L]=\{e\}$, and the intersection of C and L is a finite group. And conversely a topological group which admits such a decomposition is a completely reducible (l)-group.*

(2) *Any solvable invariant subgroup of a completely reducible (l)-group is contained in the center, which is compact.*

(3) *Any factor group of a completely reducible (l)-group is also a completely reducible (l)-group.*

REMARK. It is well known that a locally compact connected group is maximally almost periodic if and only if it is decomposable into a direct product of a compact group and a vector group.[20] On the other hand, it is clear that a vector group is not completely reducible. Hence, or from Theorem 4, we get the following proposition, by which the position of compact groups in the theory of topological groups is clarified:

[20] H. Freudenthal, loc. cit.

A locally compact connected group is a completely reducible, maximally almost periodic group if and only if it is compact.

4. Fundamental theorems.

THEOREM 5. *Let G be a locally compact connected group. If G is an (l)-group, then there exists a closed completely reducible (l)-group T in G and a closed simply connected solvable Lie invariant subgroup N such that*

$$G = TN, \qquad T \cap N = \{e\},$$

and conversely.

PROOF. Let G be an (l)-group. In virtue of the corollary to Theorem 2, we can select a sequence $\{K_\alpha\}$ of compact invariant subgroups of G such that $G_\alpha = G/K_\alpha$ is f.r. and the intersection of all K_α's coincides with the identity group. Now, because of the compactness of K_α, there corresponds $D(G_\alpha)$ to $D(G)$ in the homomorphism $G \sim G_\alpha$.

We denote by M the radical of $D(G)$. Since K_α is compact, MK_α/K_α is the radical of $D(G_\alpha)$. On the other hand we know that the radical of the commutator group of an f.r. group is simply connected.[21] Hence M cannot contain a compact subgroup other than the identity group. Therefore M is a simply connected Lie group.

Now, since $D(G)$ is an (l)-group by itself, $D(G)$ contains a compact invariant subgroup K such that the factor group $D(G)^* = D(G)/K$ is f.r. Now since $M \cap K = \{e\}$, and MK is closed, $M^* = MK/K$, which is isomorphic with M, is the radical of $D(G)^*$. Let S^* be a maximal semi-simple subgroup of $D(G)^*$. The center of S^* is finite and S^* is closed. Hence the intersection of S^* and M^* is the identity group. Thus we obtain the decomposition

$$D(G)^* = S^*M^*, \qquad S^* \cap M^* = \{e^*\},$$

where e^* denotes the identity of $D(G)^*$.

Now let S be the complete inverse image of S^* in the homomorphism $D(G) \sim D(G)^*$. Then S is obviously semi-simple and we get

$$D(G) = SM.$$

Since any solvable invariant subgroup of S is easily seen to be compact, we have

$$S \cap M = \{e\}.$$

[21] See [M.G.I, Lemma 7].

Next, since $G/D(G)$ is connected and commutative, using the structure theorem of Pontrjagin, we can assure the existence of a closed invariant subgroup H containing $D(G)$ such that G/H is compact and $H/D(G)$ is a vector group.

We denote by N the radical of H. Now we select a compact invariant subgroup K_1 of H so that H/K_1 is f.r. K_1 is obviously contained in $D(G)$. Hence we can consider K_1 as K mentioned above and retain the above notations. Let N^* be the image of N in the homomorphism $H\sim H^*=H/K$. N^* is clearly the radical of H^*. The fact that $H^*/D(G)^*$ is commutative immediately implies that any maximal semi-simple subgroup of H^* is contained in $D(G)^*$. Hence S^*, as a maximal semi-simple subgroup of $D(G)^*$, is also a maximal semi-simple subgroup in H^*. Thus we have $H^*=S^*N^*$. Hence we get that

$$H = SN.$$

Now M^* is obviously locally coincident with $D(G)^*\cap N^*$. On the other hand from the relation

$$H^*/D(G)^* \cong N^*/D(G)^* \cap N^*,$$

$N^*/D(G)^*\cap N$ is a vector group. This is impossible unless $D(G)^*\cap N^*$ is connected. Hence we have $D(G)^*\cap N^*=M^*$. Thus the simple connectedness of M^* and that of N^*/M^* imply that of N^*.

Next we shall prove that N is a simply connected Lie group isomorphic with N^*. First, since the center of S is compact, $S\cap N$ must be a compact 0-dimensional group. Hence $K\cap N$ is also compact 0-dimensional. Now suppose that $K\cap N\neq\{e\}$. Then there exists a proper open (closed) subgroup Q of $K\cap N$. Then by the relation $N/K\cap N\cong N^*$, N/Q is a proper covering group of N^*. This is clearly impossible, because N^* is simply connected. Hence $K\cap N=\{e\}$. Thus our assertion holds immediately, and moreover we get

$$S \cap N = \{e\}.$$

Next we put $G/N=G'$, $H/N=H'$. Then G' is an (L)-group, and H' is a closed connected semi-simple invariant subgroup of G' because $H'\cong S$. Hence by Corollary 2 to Theorem 1 there exists a closed invariant subgroup A_1' of G' so that $G'=H'A_1'$ and $[H', A_1'] = \{e'\}$, where e' is the identity of G'. The center of H' coincides with $H'\cap A_1'$, and is compact because H' is a semi-simple (l)-group. On the other hand, the relation $G/H\cong G'/H'\cong A_1'/H'\cap A_1'$ implies that $A_1'/H'\cap A_1'$ is compact. Hence A_1' is a compact group. Let now A' be the connected component of A_1' containing e'. Then Lemma 1

implies that $G' = H'A'$.

Next we denote by R the complete inverse image of A' in the homomorphism $G \sim G'$. R is clearly the radical of G, and we get

$$G = HR = SR.$$

Let now $G \sim (G)$ be a homomorphism of G onto an f.r. group (G) such that the kernel is compact, and let (R), (N) be the image of R, N respectively in the homomorphism. It is clear that (R) is the radical of (G) and (N) a simply connected Lie group such that $(R)/(N)$ is compact. Hence (N) is a maximal closed simply connected solvable invariant subgroup of (G). Now since (G) is f.r., (G) contains a closed completely reducible f.r. group (T) such that

$$(G) = (T)(N), \qquad (T) \cap (N) = \{(e)\}.$$

Let T be the complete inverse image of (T) in G. The fact that N contains no compact subgroup other than $\{e\}$ readily implies the following relations

$$G = TN, \qquad T \cap N = \{e\}.$$

Thus we obtain in particular the connectedness of T. Now it is easy to prove that the radical of T is compact. Hence T is completely reducible by Lemma 4. Thus our first assertion is proved.

Conversely, let G be a locally compact connected group which admits a decomposition $G = TN$, $T \cap N = \{e\}$, as above. Then, since $G/N \cong T$ and T is an (l)-group, there exists a system $\{N_\alpha\}$ of closed invariant subgroups containing N such that $\bigcap N_\alpha = N$ and G/N_α's are f.r.

Now let us denote by T_1 the set of all elements of T which are commutative with every element of N. Then T_1 is clearly a closed invariant subgroup of G such that T/T_1 is f.r. Now in virtue of Theorem 4, T/T_1 is completely reducible. Hence the relations

$$G/T_1 = T/T_1 \cdot NT_1/T_1, \qquad T/T_1 \cap NT_1/T_1 = T_1/T_1,$$
$$NT_1/T_1 \cong N,$$

imply that G/T_1 is f.r. Thus because of the relation $\bigcap N_\alpha \cap T_1 = \{e\}$ we get the result. Q.E.D.

COROLLARY. *Let G be an (l)-group and K a maximal compact subgroup of G. Then there exists a closed connected simply connected solvable Lie subgroup H of G such that*

$$G = HK = KH, \qquad K \cap H = \{e\}.$$

PROOF. In virtue of Theorem 5, it is clearly sufficient to consider the case when G is completely reducible. On the other hand, the corollary holds for f.r. groups by [M.G.II].[22] Thus by Theorem 4 we can readily complete our proof. Q.E.D.

THEOREM 6. *Let G be a locally compact connected group and R the radical of G. If R is an (l)-group and G contains a closed semi-simple (l)-group S such that*

$$G = SR,$$

then G is an (l)-group, and conversely.

PROOF. In the first place we suppose that R is an (l)-group and S is a closed semi-simple (l)-group in G such that $G=SR$. We can prove that G is an (L)-group because G/R is obviously an (L)-group.[23] Let K be a compact invariant subgroup of G such that G/K is a Lie group. Then SK/K is clearly a maximal semi-simple subgroup, and RK/K the radical, of G/K. Since they are f.r. by Theorem 2, G/K itself is also f.r. in virtue of [M.G.II]. This means that G is an (l)-group.

Conversely, let G be an (l)-group and $G=TN$ a decomposition as in Theorem 5. Let $D(T)$ be the topological commutator group of T, and A the radical of T. Then $D(T)$ is semi-simple and we get $T=D(T)A$ by Theorem 4. Now it is clear that AN coincides with the radical R. Hence we have $G=D(T)R$, and this completes our proof.

5. **Solvable (l)-groups.** Let G be a locally compact connected solvable group. Then G is an (L)-group in virtue of [K.I.]. Hence there exists a compact invariant subgroup K such that $G^*=G/K$ is a Lie group. Let now $D(G)$ be the topological commutator group of G. Then because of the compactness of K we have

$$D(G^*) = D(G)^*(= D(G)K/K).$$

Now it is well known that $D(G)^*$ is nilpotent. On the other hand K is a central invariant subgroup by [K.I.]. Hence $D(G)K$ is nilpotent, and so is $D(G)$. Thus *the topological commutator group of a locally compact connected solvable group is nilpotent.*[24]

---

[22] [M.G.II, Corollary 1 to Theorem 7].

[23] Let G be a locally compact connected group, and N a closed connected invariant subgroup of G. Suppose N and G/N are both (L)-groups. Then G is an (L)-group by itself. See [K.I.].

[24] We can prove moreover the following theorem: *The radical of the topological commutator group of a connected (L)-group is nilpotent.* This is an extension of a theorem of E. Cartan. Cf. E. Cartan, Thèse, Paris, 1894.

Now we shall prove the following lemma.

LEMMA 5. *Let G be a connected group and N a closed connected invariant subgroup. Let K be a compact subgroup of N. If N is locally compact and nilpotent, then K is central in G.*

PROOF. Let K_1 be the maximal compact invariant subgroup of N. Then K_1 is commutative and N/K_1 is a Lie group. On the other hand, we know that any compact subgroup of a connected nilpotent Lie group is a central invariant subgroup.[25] Hence we can easily prove that K is contained in K_1.

Now it is clear that K_1 is an invariant subgroup of G. Then because of the connectedness of G, K_1 must be central.[26] Hence K_1 is a central invariant subgroup of G. Q.E.D.

It is now easy to establish the following theorem.

THEOREM 7. *Let G be a locally compact connected solvable group and A a maximal compact subgroup of G. The following conditions are all necessary and sufficient for G to be an (l)-group.* (1) $G=AN$, $A\cap N=\{e\}$, *where N is a closed simply connected Lie invariant subgroup.* (2) *$D(G)$ is a simply connected Lie group.* (3) *The center of $D(G)$ is a connected simply connected Lie group.* (4) $A\cap D(G)=\{e\}$.

From Lemma 5 and Theorem 7 we have in particular the following corollary.

COROLLARY. *A locally compact connected nilpotent group is an (l)-group if and only if it is a direct product of a compact group and a connected simply connected Lie group.*

NAGOYA UNIVERSITY

[25] [M.G.II, Lemma 13].
[26] See [K.I.].

ON A CONJECTURE ON SIMPLE GROUPS

I. N. HERSTEIN

The purpose of this paper is to rephrase a conjecture about simple groups into the language of linear algebra.

Let G be a group of finite order $o(G)$. Then by Γ_p we shall mean the group ring of G over a field of characteristic p (for instance the integers modulo p). We shall denote the radical of Γ_p by N_p. If $p=0$ or $p\nmid o(G)$, then it is known that $N_p=(0)$; and if $p\mid o(G)$, $N_p\neq(0)$.

We now consider the following two assertions:

(A) *If G is a simple group of odd order, $o(G)$ is a prime.*

(B) *If G is a group of odd order $o(G)$, then for some prime p, $p\mid o(G)$, we can find a $g\in G$, $g\neq 1$, such that $g-1\in N_p$.*

The theorem which we propose to prove is:

THEOREM. (A) *is equivalent to* (B).

1. (B) implies (A).

DEFINITION. $U_p=\{g\in G \mid g-1\in N_p\}$.

LEMMA 1. *U_p is a normal p-subgroup of G.*

PROOF. (a) U_p is a subgroup of G, for if $g_1, g_2\in U_p$, then since N_p is a left-ideal of Γ_p, $g_1(g_2-1)+(g_1-1)=g_1g_2-1\in N_p$.

(b) U_p is normal, for if $g-1\in N_p$, since N_p is a two-sided ideal of Γ_p, $h(g-1)h^{-1}=hgh^{-1}-1\in N_p$ for all $h\in G$.

(c) If $g-1\in N_p$, then for some integer S, $(g-1)^{p^s}=0=g^{p^s}-1$. So if $g\in U_p$, g is of order p^s for some s. So U_p is a p-group.

COROLLARY. (B) *implies* (A).

PROOF. By (B), $U_p\neq 1$ for some $p\mid o(G)$. Hence since G is simple, and since U_p is a normal subgroup of G, $U_p=G$. Thus G is of order p^s, and G being simple, $s=1$. Hence (B) implies (A).

2. (A) implies (B).

LEMMA 2. (A) *implies that every group of odd order is solvable.*

PROOF. Let $G=G_1\supset G_2\supset\cdots\supset G_r=1$ be a composition series for G. Since the G_i/G_{i+1} are simple and of odd order, by (A) they must be of prime order; hence the lemma is proved.

Since a solvable group contains a normal p-subgroup [1, p. 25,

Presented to the Society, December 29, 1949; received by the editors May 5, 1949.

Theorem 20],[1] we immediately obtain, using Lemma 2, the following lemma.

LEMMA 3. (A) *implies that if G is of odd order, then it contains a normal p-subgroup.*

For groups of certain orders the N_p can be completely described. This is true for p-groups. If G is of order p^a, then for every $g \in G$, $g-1 \in N_p$ [2, p. 176, Theorem 1.2 or 3, p. 239]. For our case it is sufficient to use the weaker result:

LEMMA 4. *If G is of order p^a, then for some $g \neq 1$ in G, $g-1 \in N_p$.*

PROOF. Since G is of order p^a, it has a nontrivial center C. Let $g \neq 1$ be in C. Then since $g-1$ is in the center of Γ_p, and since $(g-1)^{p^a} = g^{p^a} - 1 = 1-1 = 0$, $g-1 \in N_p$.

Suppose that S is the normal p-subgroup of Lemma 3. An element of the form $g-1 \in \Gamma_p$ is in N_p if and only if for every irreducible representation ϕ of G, $\phi(g) = 1$. Clifford's theorem [4, p. 534, Theorem 1] reduces the irreducible representations of G to irreducible representations (or into ones fully reducible into irreducible components) of S. For the g of Lemma 4 in S, for every irreducible representation ϕ of S, $\phi(g) = 1$. So by Clifford's theorem, for every irreducible representation ϕ of G, $\phi(g) = 1$. Thus $g-1 \in N_p$. And so we have shown the following lemma.

LEMMA 5. (A) *implies* (B).

BIBLIOGRAPHY

1. A. Speiser, *Gruppen der endlichen Ordnung.*

2. S. A. Jennings, *The group ring of a p-group over a modular field*, Trans. Amer. Math. Soc. vol. 50 (1941).

3. Lombardo-Radice, *Intorno alle algebre legato ai gruppi di ordine finito*, Rendiconti del Seminario Matematico della università di Roma vol. 3 (1939).

4. A. H. Clifford, *Representations induced in an invariant subgroup*, Ann. of Math. vol. 38 (1937).

UNIVERSITY OF KANSAS

[1] Numbers in brackets refer to the bibliography at the end of the paper.

ON A SEMI-GROUP OF SUBSETS OF A LINEAR SPACE

R. E. FULLERTON

A family $\mathcal{F}$ of subsets of a space is defined to be a semi-group of subsets provided that the intersection of any two subsets of $\mathcal{F}$ is again a subset of $\mathcal{F}$. In giving a characterization of a Banach space of continuous functions in terms of the geometry of the space, Clarkson[1] considered a cone $\mathfrak{C}$ which had the property that the family of all of its translates formed a semi-group. This note is concerned with an investigation of the structure of a subset S of a linear space $\mathfrak{X}$ such that the family of all translates of S form a semi-group, that is, for any points x, y of $\mathfrak{X}$ there exists a $z\in\mathfrak{X}$ such that $(x+S)\cap(y+S)=z+S$. It will be shown that under certain rather weak restrictions, S must be a convex cone.

Let $\mathfrak{X}$ be an abstract linear space. The only topology which will be assumed for $\mathfrak{X}$ is that which arises from the topology of the straight lines of $\mathfrak{X}$, that is, if $u_n=\alpha_n x+\beta_n y$, $x\neq y$, $\alpha_n+\beta_n=1$, $\alpha_n\to\alpha$, $\beta_n\to\beta$, then $u_n\to\alpha x+\beta y$. This will be called the linear topology of $\mathfrak{X}$. We shall assume that all straight lines of $\mathfrak{X}$ are complete in this topology. A point x is an extreme point of a set S if $x\in S$ and there exists no line segment with end points in S which contains x in its interior. The theorem to be proved is the following:

THEOREM. *Let S be a subset of $\mathfrak{X}$ which is closed in the linear topology and which has at least one extreme point. Let the family of all translates of S form a semi-group. Then S is a convex cone, that is, S is convex and there exists a point $v\in S$ such that*

$$S = \underset{v}{E}\,[y = v + \lambda x,\ x \in S,\ \lambda \geq 0].$$

The theorem is proved by four lemmas.

LEMMA 1. *If $x+S=S$, then $x=\theta$.*

PROOF. It is evident that any translate of S can be used initially instead of S. Thus we may assume that S contains θ without loss of generality. If $x+S=S$, $x\in S$. Also $S=S-x$ and $-x\in S$. If $y\in S$, $x+y\in S$ and $-x+y\in S$. However, if $x\neq -x$ this shows that y is the midpoint of a line segment joining two points of S and con-

Presented to the Society, December 30, 1948; received by the editors January 4, 1949 and, in revised form, June 29, 1949.

[1] J. A. Clarkson, *A characterization of C spaces*, Ann. of Math. vol. 48 (1947) pp. 845–850.

tradicts the hypothesis that S has an extreme point.

LEMMA 2. *S contains only one extreme point.*

PROOF. Suppose that S contained at least two extreme points. It may be assumed that one is the point θ. Let the second be $v \neq \theta$. Consider the translation $S+v$. Since θ and v are both extreme points, $2v \notin S$ and $S+v$ does not contain S. Hence $(S+v) \cap S = y+S$ where $y \neq \theta$ and y is an extreme point of $y+S$. Also

$$\{v + [(S+v) \cap S]\} \cap S$$
$$= (v + y + S) \cap S \supset (v + y + S) \cap (y + S)$$
$$= y + (y + S) = 2y + S.$$

This shows that v, $y+v$, and $2y+v$ are in S and $y+S$ contrary to the conclusion that y is an extreme point of $y+S$.

LEMMA 3. *If v is the extreme point of S, S is a cone with vertex v.*

PROOF. It may be assumed that $v=\theta$. It must be shown that if $x \in S$, then $\lambda x \in S$ for all $\lambda \geqq 0$. Since λx is a complete ray, if there exist values of λ for which $\lambda x \notin S$, there must exist a smallest value λ_0 such that $\lambda_0 x \in S$ because of the linear closure of S. $(\lambda_0 x + S) \cap S = y + S$ where y is the extreme point of $y+S$. However, since $\lambda_0 x \in (\lambda_0 x + S) \cap S$, $\lambda_0 x$ is an extreme point of $y+S$ and since there is only one extreme point of the set, $y = \lambda_0 x$ by Lemma 1. Then $\lambda_0 x + S \subset S$ and $2\lambda_0 x \in S$. By induction it can be seen that if $\lambda_0 x \in S$, then $n\lambda_0 x \in S$ for any positive integer n. Also there exists a $z \in S$ such that $(2^{-1}\lambda_0 x + S) \cap S = z + S$. This implies the existence of a point $u \in S$ such that $z = u + 2^{-1}\lambda_0 x$. Also, by the first part of the proof, $u + \lambda_0 x \in S$. Since z is the extreme point of $z+S$, $u + (n/2)\lambda_0 x \in z + S \subset S$ for every integer n by the same argument as was used in the first part of the proof. Then $S = S - z$ contains all points of the form $u + (n/2)\lambda_0 x - (u + 2^{-1}\lambda_0 x)$. In particular, S contains the point $2^{-1}\lambda_0 x$, contrary to the assumption that λ_0 was the smallest nonzero value of λ for which $\lambda x \in S$. This shows that if $x \in S$, $\lambda x \in S$ for all $\lambda \geqq 0$ and S is a cone with vertex θ.

It is to be noted that if θ is the vertex, the above proof shows that if $x \in S$, then $\lambda x \in S$ for all non-negative λ and $\lambda x + S \subset S$. Hence, if $y \in S$, $y + \lambda x \in S$ for all $\lambda \geqq 0$. This shows that if S contains a ray through any point, it contains a parallel ray through every point of the set.

LEMMA 4. *S is convex.*

PROOF. It must be shown that if x and y are in S, then $\alpha x+\beta y\in S$ for $\alpha, \beta\geqq 0, \alpha+\beta=1$. If S has vertex θ and if $y=\lambda x$, this is true by Lemma 3. If $y\neq\lambda x$, then the rays λx, νy, $\lambda\geqq 0$, $\nu\geqq 0$, are distinct and are in S. If for any point $\alpha x+\beta y$, $\alpha, \beta>0$, $\alpha+\beta=1$, there exists a $\mu_0\neq 0$ such that $\mu_0(\alpha x+\beta y)\in S$, then $\mu(\alpha x+\beta y)\in S$ for all $\mu\geqq 0$ by Lemma 3. S contains the ray $y+\lambda x$ where $\lambda\geqq 0$. If the existence of positive numbers μ and λ can be demonstrated such that $\mu(\alpha x+\beta y)=y+\lambda x$, the lemma is proved. If this is so, $(\mu\alpha-\lambda)x+(\mu\beta-1)y=\theta$ and if $\mu=1/\beta$, $\lambda=\alpha/\beta$, the rays $\mu(\alpha x+\beta y)$ and $y+\lambda x$ have a common point. Since all points of the form $y+\lambda x$ are in S, $(1/\beta)(\alpha x+\beta y)\in S$ and hence $\mu(\alpha x+\beta y)\in S$ for all $\mu\geqq 0$, in particular for $\mu=1$. Hence S is convex.

By making use of this theorem it is possible to weaken the hypotheses of Clarkson's theorem slightly and to state it in the following form.

THEOREM. *A necessary and sufficient condition that a Banach space be equivalent to the space of all continuous functions over a bicompact Hausdorff space H is that there exist a closed set S with an extreme point, containing θ in its interior, such that all translates of S form a semi-group and such that the unit sphere in the space is the set $S\cap(-S)$.*

UNIVERSITY OF WISCONSIN

NOTE ON THE ARITHMETIC OF BILINEAR TRANSFORMATIONS

DONALD M. ADELMAN

1. Introduction. Since the time of Lucas many papers have appeared investigating the properties of sequences of rational integers satisfying linear recurrence relations. Very little,[1] however, has been done in the nonlinear cases, although the generalization to variable coefficients seems to have awakened more interest than the generalization to relations of higher degree but with fixed coefficients. The present paper deals with one of the simplest problems of the latter type.

Let a sequence be defined by the value x_1 and by the relation:

$$x_{n+1} = (ax_n + b)/(cx_n + d); \tag{1}$$

$ad-bc\neq 0$; a, b, c, d, rational integers. Under what conditions are all the x_i integral? That such integral sequences exist may be shown by two examples. First, if $x_{n+1}=x_n/(x_n-1)$, $x_1=2$ generates the sequence 2, 2, 2, $\cdots$; second, if $x_{n+1}=(178x_n-1492)/(7x_n+2)$, $x_1=4$ generates the sequence 4, -26, 34, 19, 14, 10, 4, $\cdots$. If we write $y_n=cx_n+d$, $e=a+d$, $m=bc-ad$, (1) becomes:

$$y_{n+1} = e + m/y_n \tag{2}$$

where $\{y_i\}$ is integral if $\{x_i\}$ is.

THEOREM 1. *Every sequence of integers satisfying* (1) *or* (2) *is periodic.*

PROOF. In (2) every y_n must be one of the finitely many divisors of m. The periodicity of $\{y_i\}$ entails that of $\{x_i\}$. Moreover, the period begins with the first term of the sequence, since each value of x_n permits only one value for x_{n-1}.

2. Classification by fixed points. The possible sequences satisfying (2) are related to the roots, r_1 and r_2, of $r^2-er-m=0$.

THEOREM 2. *If the roots are real and of unequal magnitude there are no sequences satisfying* (2) *or just two such sequences, each of period one. In the latter case r_1 and r_2 are integral and generate the two sequences.*

Received by the editors May 7, 1949.

[1] An exception is Morgan Ward's recent memoir on "elliptic" divisibility sequences in vol. 70 of the American Journal of Mathematics.

PROOF. Suppose $|r_1| < |r_2|$ and write (2) as:

$$\frac{y_{n+1} - r_2}{y_{n+1} - r_1} = \frac{r_1}{r_2}\,\frac{y_n - r_2}{y_n - r_1} = \left(\frac{r_1}{r_2}\right)^n \frac{y_1 - r_2}{y_1 - r_1}. \tag{3}$$

$y_1 = r_1$ only if r_1 is integral and the sequence is $r_1, r_1, r_1, \cdots$. Otherwise $\lim_{n\to\infty}(r_1/r_2)^n = 0$, and $\lim_{n\to\infty}(y_{n+1} - r_2)/(y_{n+1} - r_1) = 0$, which is only possible if $y_{n+1} = r_2$.

If the roots are real and equal except for sign, $e = 0$ so that we may take y_1 as any divisor of m and the sequence becomes:

$$y_1,\ m/y_1,\ y_1,\ m/y_1, \cdots.$$

If $r_1 = r_2 = r$, (3) is of no use and instead we define a sequence, $p_0, p_1, p_2, \cdots$ by $p_0 = 1$, $p_1 = y_1$, $p_{n+2} = ep_{n+1} + mp_n$. Then:

LEMMA 1. $y_n = p_n/p_{n-1}$.

PROOF. Set $p_n/p_{n-1} = q_n$. Then $q_1 = y_1$, and the recurrence relation becomes:

$$q_{n+2}q_{n+1}p_n = eq_{n+1}p_n + mp_n. \tag{4}$$

Now, from the theory of second order recurrence sequences we know that the general solution for p_n is $p_n = (An + B)r^n$, where A and B are not both zero since $p_0 = 1$. Evidently $p_n = 0$ for at most one value $n = t$. For all values of n except t, (4) then reduces to

$$q_{n+2} = e + \frac{m}{q_{n+1}},$$

which is the same relation satisfied by the y_i without exception. The q-sequence and the y-sequence therefore agree in their first $t+1$ terms. Also, in both, the $t+2$nd term is equal to e, in the y-sequence because y_{t+1} is infinite, in the q-sequence by setting $p_t = 0$ in the second order recurrence relation and using the definition of q_{t+2}. Therefore the two sequences agree in all their terms and the lemma is proved.

Thus it is required that $((An+B)/(A(n-1)+B))r = y_n$ should be integral for all values of n. But, if $A = 0$, $y_n = r$, and otherwise $\lim_{n\to\infty} y_n = r$, so that again the only possible sequence occurs when r is integral. This gives:

THEOREM 3. *If $r_1 = r_2 = r$, there is just one sequence satisfying* (2) *or none according as r is or is not an integer. In the former case the sequence is $r, r, r, \cdots$.*

Inasmuch as we have already seen an example of a sequence of

period six, we may expect the case of complex roots to be more difficult. For real roots, of course, it is hardly necessary to use (2) since (1) and the value of e give all the required information.

3. **Complex roots.** Again we consider the linear recurrence relation:

$$p_{n+1} = ep_n + mp_{n-1}, \tag{5}$$

but this time define a sequence $\{p_i\}$ by the initial values $p_0=0$, $p_1=1$. This is called the principal sequence, and its first seven terms are $0, 1, e, e^2+m, e^3+2em, e^4+3e^2m+m^2, e^5+4e^3m+3em^2$. Using this sequence, we obtain the following result:

THEOREM 4. *If* $\{y_i\}$ *satisfies* (2) *and* $\{p_i\}$ *is the principal solution of* (5), *then*

$$y_n = \frac{p_n y_1 + mp_{n-1}}{p_{n-1}y_1 + mp_{n-2}}.$$

The proof is by induction.

THEOREM 5. *The sequence* $\{y_i\}$ *satisfying* (2) *has a period* w *if and only if* $p_w=0$.

PROOF. Suppose $p_w=0$. Then by (5), $p_{w+1}=mp_{w-1}$ and from Theorem 4, $y_{w+1}=y_1$. On the other hand, suppose $y_{w+1}=y_1$. Then:

$$p_w y_1^2 + mp_{w-1}y_1 = p_{w+1}y_1 + mp_w,$$
$$p_w y_1^2 - (p_{w+1} - mp_{w-1})y_1 - mp_w = 0 = p_w y_1^2 - ep_w y_1 - mp_w,$$
$$p_w(y_1^2 - ey_1 - m) = 0,$$

and since the factor in parentheses cannot be zero, $p_w=0$.

Of course the period of $\{y_i\}$ is the smallest value of $w>0$ for which $p_w=0$.

LEMMA 2.[2] *If* $\{p_i\}$ *is the principal solution of* (5), *for* $n>0$ p_n *is a homogeneous polynomial of degree* $[(n-1)/2]$ *in the variables* e^2 *and* m *if* n *is odd, and* e *times such a polynomial if* n *is even. In both cases all the coefficients are integers and the coefficient of the highest power of* e *is unity.*

---

[2] The proof beginning here is not the one that most readily suggests itself. If p_n is expressed explicitly in terms of the scalar roots, namely, as $p_n=(r_1^n-r_2^n)/(r_1-r_2)$, the problem of finding its zeros reduces to that of determining the exponent of an element in a quadratic field. But the proof given above relies on simpler and somewhat more relevant ideas.

The proof is by induction.

It follows from Theorem 5 that if (2) is satisfied by some integral sequence $\{y_i\}$ with period w, the values of e and m from (2) make the homogeneous polynomial of Lemma 2 zero. Consequently e^2/m is the root of a polynomial with integral coefficients and leading coefficient unity. Therefore, since the root is rational it must be integral and we have:

LEMMA 3. *$p_w=0$ only if $e^2+km=0$ for some integer, k.*

Not all values of k correspond to a w however, as is seen from:

THEOREM 6. *If r_1 and r_2 are complex, and if $\{y_i\}$ satisfies (2) with period w, then $e^2+km=0$ for one of four values of k, $k=0,1,2,$ or 3, corresponding respectively to $w=2, 3, 4,$ or 6.*

PROOF. Since the roots are complex, $e^2+4m<0$ and $m<0$. Therefore $e^2+km>0$ for $k<0$, and $e^2+km<0$ for $k\geqq 4$. Only four values of k remain, and we obtain the corresponding periods from the factorization of the first seven terms of the sequence $\{p_i\}$.

4. **The complete period.** We now know that any sequence $\{y_i\}$ of integers satisfying (2) must be periodic, and conditions on e and m in (2)—or on a, b, c, and d in (1)—for periodicity. But these conditions do not guarantee that for an arbitrary value of y_1, or even for any value of y_1, the sequences $\{x_i\}$ and $\{y_i\}$ should be entirely integral. We establish certain preliminary theorems.

THOREM 7. *If w is the period of a sequence $\{y_i\}$ of integers satisfying (2), then $(y_1y_2\cdots y_w)^2=(-m)^w$.*

PROOF. $r_i=e+m/r_i$ for $i=1$ and 2. Subtracting both these equations from (2) and multiplying the results, we obtain:

$$(y_{n+1}-r_1)(y_{n+1}-r_2)=\frac{m^2}{r_1r_2y_n^2}(y_n-r_1)(y_n-r_2)$$

or, if $Q(y)=y^2-ey-m$,

$$Q(y_{n+1})=\frac{-m}{y_n^2}Q(y_n).$$

Multiplying these equations over a complete period we obtain the result, since $Q(y)\neq 0$ for y real.

COROLLARY 1. *If an integer s divides each term of $\{y_i\}$, then $s^2\mid m$ and $s\mid e$.*

PROOF. From Theorem 7, $s^{2w} \mid (-m)^w$ or $s^2 \mid m$. Then, from Theorem 6, $s \mid e$.

COROLLARY 2. *If* $\{x_i\}$ *is a sequence of integers satisfying* (1) *and if* $s=(c, d)$, *then* $s \mid a$ *and* $s \mid b$. *Accordingly we may assume* $(c, d)=1$.

PROOF. Every term of the sequence $\{y_i\}$ derived from $\{x_i\}$ is divisble by s. By Corollary 1, $s \mid e=a+d$, and therefore $s \mid a$. Writing (1) as

$$x_{n+1}(cx_n + d) = ax_n + b,$$

we have $s \mid b$.

DEFINITION. *Let* $R(g, h)$ *be the greatest divisor of* g *prime to* h. Then, for example, $R(6, 9)=2$; $R(9, 6)=1$.

THEOREM 8. *An integral sequence* $\{x_i\}$ *satisfying* (1) *corresponds to an integral sequence* $\{y_i\}$ *satisfying* (2) *if and only if* c *is a divisor of* $R(y_1-y_2, y_1)$.

PROOF. If we begin with the sequence $\{x_i\}$, then $y_1=cx_1+d$, $y_2=cx_2+d$, and $c \mid (y_1-y_2)$. By Corollary 2 we may assume that c is prime to d, consequently to each y_i, and in particular to y_1.

Start, however, from a sequence $\{y_i\}$ and let c be any divisor of $R(y_1-y_2, y_1)$. Choose an arbitrary integer x_1 and define d by $y_1=cx_1+d$. Then, since $(c, y_1)=1$, $(c, d)=1$. Next, define a by means of $e=a+d$. Now, since $c \mid (y_1-y_2)$, $y_1 \equiv y_2 \equiv d \pmod{c}$, and writing (2) as

$$y_{n+1}y_n = ey_n + m, \tag{6}$$

we have $d^2 \equiv ed+m \pmod{c}$ by taking $n=1$, or, $d(d-e) \equiv m \equiv -ad \pmod{c}$, using the definition of a. So, $(m+ad)/c$ is an integer, say b. We thus have found integers, a, b, c, and d, such that, in (2), $e=a+d$ and $m=bc-ad$. It remains to show that all the x_i are integral. Obviously, this will be true if and only if $y_i \equiv d \pmod{c}$ for each i. We already know this for $i=1$, and from (6), if $y_n \equiv d \pmod{c}$,

$$y_{n+1}d \equiv (a + d)d - ad \pmod{c} \equiv d^2$$

or, since $(c, d)=1$, $y_{n+1} \equiv d \pmod{c}$.

COROLLARY. *For any given sequence* $\{y_i\}$ *of integers satisfying* (2), $R(y_n-y_{n+1}, y_n)$ *is a constant.*

PROOF. Since $\{y_i\}$ is periodic, it does not matter which term we consider the first in defining a value of c as in Theorem 8. In particular, the largest possible value of c cannot change.

5. **Parametric solution.** Theorem 8 completely elucidates the connection between sequences satisfying (2) and sequences satisfying (1). We shall now give a parametric solution to the problem of finding all possible sequences satisfying (2). Evidently this is hardly necessary when r_1 and r_2 are real or when $e=0$, so that by Theorem 6, $k=-e^2/m$ must assume one of the three values 1, 2, or 3. Introducing an integral parameter t, we must have, according to the value of k, $e=t$ and $m=-t^2$, or $e=2t$ and $m=-2t^2$, or $e=3t$ and $m=-3t^2$. This gives the first part of the following theorem.

THEOREM 9. *If $e^2+4m<0$ and $e\neq 0$, the transformation* (2) *must take one of three forms*:

$$y_{n+1} = k(t - t^2/y_n), \qquad k = 1, 2, \text{ or } 3,$$

where t is an integer. In addition, there must exist relatively prime integers, u and v, such that t is divisible by $uv(u-v)$ if $k=1$, by either $uv(u-v)(u-2v)$ or $uv(u-v)(u-2v)/4$ according as u is odd or even if $k=2$, or by either $uv(u-v)(2u-3v)(u-2v)(u-3v)$ or $uv(u-v)(2u-3v)(u-2v)(u-3v)/27$ according as u is not or is divisible by 3. In all cases $y_1=ut/v$.

PROOF. The proof is much the same for any value of k and will be given for $k=2$. If a sequence $\{y_i\}$ of integers satisfies $y_{n+1}=2t-2t^2/y_n$, we may certainly find relatively prime integers, u and v, such that $y_1=ut/v$. Then y_2, y_3, and y_4 may be calculated as:

$$2t\frac{u-v}{u}, \qquad t\frac{u-2v}{u-v}, \quad \text{and} \quad 2t\frac{-v}{u-2v}.$$

It is easily seen that if the fraction u/v is in its lowest terms so are $(u-v)/u$, $(u-2v)/(u-v)$, and $v/(u-2v)$. Consequently, $v \mid t$, $u \mid 2t$, $(u-v) \mid t$, and $(u-2v) \mid 2t$. If u is odd, u, v, $u-v$, and $u-2v$ have no factor in common since $(u, v)=1$. If u is even, $(u, u-2v)=2$. The theorem follows.

EXAMPLE. Take $u=4$, $v=3$. Then t must be divisible by $4\cdot 3\cdot 1\cdot(-2)/4=-6$. A possible value for t is $t=12$ which gives $y_1=16$, the transformation being: $y_{n+1}=24-288/y_n$, and the sequence: 16, 6, -24, 36, 16, $\cdots$, $R(y_1-y_2, y_1)=R(10, 16)=5$. Thus, to construct a sequence satisfying (1) we may take $c=5$, and perhaps—the choice is arbitrary—$x_1=3$; then $d=1$, $a=e-d=23$, $b=(m+ad)/c=-53$. The transformation is therefore $x_{n+1}=(23x_n-53)/(5x_n+1)$, and the sequence: 3, 1, -5, 7, 3, $\cdots$.

UNIVERSITY OF CONNECTICUT

MINIMAL IDENTITIES FOR ALGEBRAS

A. S. AMITSUR AND J. LEVITZKI

1. **Introduction.** Let F be the underlying field of an algebra A and $x_1, x_2, \cdots, x_m$ a set of indeterminates. Consider a polynomial $f(x_1, \cdots, x_m)$ over F, that is, an element of the free algebra $F(x_1, \cdots, x_m)$ generated by the indeterminates x_i over the field F. If this polynomial is not identically zero and if the equation

$$f(x_1, \cdots, x_m) = 0 \tag{1}$$

is satisfied by all elements of A, then we say (see [1][1]) that the polynomial identity (1) holds in A. Such identities are satisfied, for example, by each algebra of finite dimensionality.[2] An identity of minimal degree will be called a *minimal identity*, and the corresponding polynomial a minimal polynomial.[3] In a previous communication (see [4]) it has been shown by one of the present authors that the degree of the minimal identities for the total matric-algebra A_n of all n by n matrices is at least $2n$. In the present note we show (§2) that *this degree is exactly* $2n$, and that we may choose as a polynomial identity one having the form

$$S(x_1, \cdots, x_{2n}) = \sum_{(i)} \pm x_{i_1} x_{i_2} \cdots x_{i_{2n}} \tag{2}$$

where the sum on the right ranges over all permutations of $2n$ letters and the sign is positive for even permutations and negative for odd permutations. We shall use for (2) occasionally also the shorter notation $S_{2n}(x)$. Polynomials of type (2) were previously studied in the same context by I. Kaplansky [2] and F. W. Levi [3], and we adopt for these polynomials and the corresponding identities the terms *standard polynomials*, respectively *standard identities*, suggested by Kaplansky.

In §3 we deal with the problem of determining all minimal polynomials of A_n. It turns out that all minimal polynomials are homogeneous and linear in each of the indeterminates. Moreover, if $x_1, x_2, \cdots, x_k$ is a given set of indeterminates, then minimal polynomials depending on the x's only may be constructed if and only if

Received by the editors May 8, 1949.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

[2] For further details compare [1] and [4].

[3] Minimal polynomials in the ordinary sense will not be referred to in this paper, so that no confusion will arise.

$k \geqq 2n$ and the set of all minimal polynomials in the x's forms (if supplemented by zero) a module over the underlying field whose dimensionality is $C_{k,2n}$. As a basis for this module we may choose the $C_{k,2n}$ standard polynomials $S(x_{i_1}, \cdots, x_{i_{2n}})$, where $(i_1, \cdots, i_{2n})$ is an arbitrary combination of $2n$ letters out of k. For $k=2n$ we have in particular the following uniqueness theorem: The minimal polynomial depending on no more than $2n$ indeterminates is uniquely determined (up to a numerical factor) as the standard polynomial $S_{2n}(x)$. The case $n \leqq 2$, $F=$prime field of characteristic 2 requires a special treatment.

In §4 we apply our results to general algebras of finite dimensionality.

2. The main theorem. We begin with a few remarks concerning some simple properties of the standard polynomials which will be needed in the proof of the main theorem.

REMARK 1. Since the standard polynomial is homogeneous and linear in each of its indeterminates, we have

$$S(x_1, \cdots, x_i, \alpha u + \beta v, x_{i+2}, \cdots) = \alpha S(x_1, \cdots, x_i, u, \cdots, x_{2n}) + \beta S(x_1, \cdots, x_i, v, \cdots, x_{2n}).$$

REMARK 2. The standard polynomial vanishes if two of the indeterminates are identified.

REMARK 3. If $(i_1, \cdots, i_m)$ is any permutation of m letters, then $S(x_{i_1}, \cdots, x_{i_m}) = \pm S(x_1, \cdots, x_m)$.

REMARK 4. Denote by $(i_1, \cdots, i_m)$ any permutation of m letters, and by S' the sum of all terms in $S_m(x)$ having in common the left factor $x_{i_1}x_{i_2} \cdots x_{i_r}$, where $r<m$, then $S' = \pm x_{i_1}x_{i_2} \cdots x_{i_r} S(x_{i_{r+1}}, \cdots, x_{i_m})$.

REMARK 5. Denote by $a_1, a_2, \cdots, a_m$ a set of m matrices of order m, such that the rth row and the rth column of each of these matrices consists of zeros. Denote by a_i' the matrix derived from a_i by deleting the rth row and the rth column. If $S(a_1, \cdots, a_m)=0$, then also $S(a_1', \cdots, a_m')=0$ and conversely. This is a consequence of the fact that the correspondence $a \to a'$ determines an isomorphism between the rings generated by the a_i and the a_i' respectively.

REMARK 6. If an algebra satisfies an identity $S_m(x)=0$, then it satisfies also each identity $S_k(x)=0$ with $k>m$.

LEMMA 1. *If for an odd positive integer r we put $y=x_{i+1} \cdots x_{i+r}$ and if S' denotes the sum of all terms of $S_m(x)$ containing the common factor y, then*

$$S' = S(x_1, \cdots, x_i, y, x_{i+r+1}, \cdots, x_m). \tag{3}$$

PROOF. The right member of (3) is a sum of terms with coefficients ± 1 each being a product of $m-r+1$ factors. If p and q are any two such terms, then let π denote the permutation transforming p into q. If y is replaced by the product $x_{i+1} \cdots x_{i+r}$, then p and q become products of m factors each, and the permutation π induces a permutation π' of these m factors. Relation (3) will be proved if we show that these two permutations are of the same type, that is, that π and π' are even or odd simultaneously. It is evidently sufficient to show that if π is a transposition, then π' is odd. Now this is evident in case π does not involve the factor y, since in this case π' is also a transposition. Suppose now that π permutes y and x_j, and denote by k the number of factors separating y from x_j. It follows easily that the permutation π' may be resolved into $k(r+1)+r$ transpositions of adjacent factors. Since r is odd, also $k(r+1)+r$ is odd, that is, the permutation π' is odd, q.e.d.

We now turn to the total matric-algebra A_n and introduce the n^2 matric-units e_{ij}, which are subject to the relations

$$e_{ij}e_{j'k} = \begin{cases} e_{ik} & j = j', \\ 0 & j \neq j', \end{cases} \tag{4}$$

and consider sets of $2n$ matric-units

$$e_{i_1 j_1}, e_{i_2 j_2}, \cdots, e_{i_{2n} j_{2n}}. \tag{5}$$

Our chief concern will be to investigate the quantity

$$S(e_{i_1 j_1}, e_{i_2 j_2}, \cdots, e_{i_{2n} j_{2n}}) \tag{6}$$

and to show that this quantity vanishes. It will be convenient to introduce a function $f(u)$ defined for $1 \leqq u \leqq n$ which denotes the number of occurrences of u among the $4n$ subscripts in (5). Evidently

$$0 \leqq f(u) \leqq 4n, \qquad \sum_{u=1}^{n} f(u) = 4n. \tag{7}$$

We now prove the following lemma.

LEMMA 2. *Suppose that the identity* $S_{2n-2}(x) = 0$ *holds for any system of* $2n-2$ *units in* A_{n-1}. *Suppose further that at least one of the units* (5) *is idempotent, and that for this unit* e_{uu} *we have* $f(u) \leqq 4$. *Then the quantity* (6) *is equal to* 0.

PROOF. In case any two of the units (5) are equal, this is true by Remark 2, and thus we may assume that all units in (5) are dif-

ferent from each other. Since $2 \leqq f(u) \leqq 4$, we have to consider three cases. The case $f(u)=2$ is trivial, since here the idempotent e_{uu} is the only unit with a subscript equal to u, and hence by (4) each term in (6) vanishes. In case $f(u)=3$, we have aside from e_{uu} a unit of the form e_{uk}(resp. e_{ku}) while the remaining $2n-2$ units have no subscripts equal to u. By (4) it follows that the only terms which are possibly not equal to 0 are those with the left (resp. right) factor $e_{uu}e_{uk}$ (resp. $e_{ku}e_{uu}$). If therefore we denote the remaining $2n-2$ factors by $e^{(1)}, e^{(2)}, \cdots, e^{(2n-2)}$, it follows in view of Remark 4 that (6) is equal to $\pm e_{uu}e_{uk}S(e^{(1)}, \cdots, e^{(2n-2)})$ (resp. $=\pm S(e^{(1)}, \cdots, e^{(2n-2)})e_{ku}e_{uu}$). Since none of the units $e^{(i)}$ has a subscript equal to u, it follows by the assumption of the lemma in view of Remark 5 that $S(e^{(1)}, \cdots, e^{(2n-2)})=0$, and hence also (6) vanishes. It remains to deal with the case $f(u)=4$. Here we have exactly three units having a subscript equal to u. If these units are of the form e_{uu}, e_{uk}, e_{uj} or of the form e_{uu}, e_{ku}, e_{ju}, then by (4) each term in (6) vanishes, that is, (6) is equal to 0. We may therefore assume that the three units with the subscript u are of the form e_{uu}, e_{uk}, e_{ju}. Denote the remaining $2n-3$ units by $e^{(1)}, \cdots, e^{(2n-3)}$ and define the three partial sums S', S'', S''' of (6) as follows: S' is the sum of all terms with the left factor $e_{uu}e_{uk}$ and the right factor e_{ju}; S'' is the sum of all terms with the right factor $e_{ju}e_{uu}$ and the left factor e_{uk}; S''' is the sum of all terms with the common factor $e_{ju}e_{uu}e_{uk}=e_{jk}$. Since by (4) all other terms in (6) vanish, we have

$$S(e_{i_1 j_1}, \cdots, e_{i_{2n} j_{2n}}) = S' + S'' + S'''. \tag{8}$$

Each nonzero term of S' is of the form $e_{uu}e_{uk}ae_{ju}=\pm e_{uu}$, where a is a product of $2n-3$ units. To this term uniquely corresponds a term with an opposite sign belonging to S'', namely $\mp e_{uk}ae_{ju}e_{uu}=\mp e_{uu}$. The signs of the corresponding terms are indeed opposite since the number of factors is even $(=2n)$ and hence the cyclic permutation which transforms the corresponding terms in one another is odd. Hence it follows that $S'+S''=0$. As to the partial sum S''' we have in view of Remark 3 and Lemma 1

$$S''' = \pm S(e_{jk}, e^{(1)}, e^{(2)}, \cdots, e^{(2n-3)}). \tag{9}$$

Since the subscripts of the $2n-2$ units participating in (9) are different from u, it follows by the assumption of the lemma in view of Remark 5 that $S'''=0$. This completes the proof of the lemma.

LEMMA 3. *Suppose that $S_{2n-2}(x)=0$ is valid for any system of $2n-2$ units in A_{n-1}. Suppose further that* (6) *vanishes whenever the number of*

the idempotents among the units (5) *is not less than* $r+1$, *where* $r \geqq 0$. *Then* (6) *also vanishes whenever the number of idempotents in* (5) *is equal to* r.

PROOF. As in the proof of Lemma 2 we easily dispense with the case that any of the units are equal, and in case $r>0$ may assume without loss of generality that the idempotents are $e_{11}, e_{22}, \cdots, e_{rr}$. If for some subscript u such that $1 \leqq u \leqq r$ we have $f(u) \leqq 4$, then (6) vanishes by Lemma 2. We may therefore restrict our attention to the case that $f(u) \geqq 5$ for $u \leqq r$ and consider the following two cases:

Case I. There exists in (5) a unit e_{ik} such that $i>r$, $k>r$. This holds, for example, for each unit in case $r=0$. Consider the matrix $c=e_{ii}+e_{ik}$. This matrix is idempotent of rank 1 and the elements

$$e_{11}, \cdots, e_{rr}, c \tag{10}$$

form an orthogonal system of $r+1$ idempotents,[4] each having the rank 1.

Case II. For each unit e_{ik} and e_{ki} with $i>r$ we have $k \leqq r$. This implies, for example, that $r \geqq 1$. Since $f(1) \geqq 5$, it follows that for some i such that $i>r$ we must have $f(i) \leqq 3$. We now reduce this to the case where we may assume that for some i we even have $f(i) \leqq 2$. To this end first note that it follows by (4) that if some term in (6) is not equal to 0, then either each $f(u)$ is even, or $f(u)$ is even for $2n-2$ indices and odd for 2 indices. Suppose now that $f(u) \geqq 3$ for all u's and $f(i)=3$ for some i. Then either each term in (6) vanishes and our lemma is proved, or else we are confronted by the following two possibilities:

(1) $f(i)=f(j)=3$ for a pair of indices $i \neq j$, and $f(u) \geqq 4$ for $u \neq i, j$.

(2) $f(i)=3$, and $f(u) \geqq 4$ for $u \neq i$.

We may assume in the first of these two cases that $f(1)$ is even, and since $f(1) \geqq 5$, we must have $f(1) \geqq 6$, and thus it follows in both cases in view of (7) that we must have $r=1$. This implies that each unit is either of the form e_{s1} or e_{1s}, which leaves us with at most $2n-1$ units instead of with $2n$. This contradiction shows that in Case II we may indeed assume that for some i we have

$$i > r, \qquad f(i) \leqq 2. \tag{11}$$

If for this i we have $f(i)=0$, our lemma is true by Remark 5. Thus Case II is reduced to the case where we have a unit e_{ik} (or e_{ki}) so that (11) holds and $k \leqq r$. We assume without loss of generality that $k=1$. By putting $c=e_{ii}+e_{i1}$ (respectively $c=e_{ii}+e_{1i}$) and $d=e_{11}-e_{i1}$ (respectively $d=e_{11}-e_{1i}$), we get the following orthogonal set of $r+1$

[4] In case $r=0$, this system reduces to the single idempotent c.

idempotents of rank 1

$$d, e_{22}, \cdots, e_{rr}, c. \tag{12}$$

In view of the properties of systems (10) and (12) it follows that there exists a regular matrix a with coefficients in F which transforms (10), respectively (12), into the system of $r+1$ idempotents

$$e_{11}, \cdots, e_{rr}, e_{r+1\,r+1}. \tag{13}$$

In both Cases I and II we may write our units in the form $e_{11}, \cdots, e_{rr}, e_{ik}, e^{(1)}, \cdots, e^{(2n-r-1)}$ (in Case II we must eventually replace e_{ik} by e_{ki}) and we have in view of Remark 3

$$\pm S(e_{i_1 j_1}, \cdots, e_{i_{2n} j_{2n}}) = S(e_{11}, \cdots, e_{rr}, e_{ik}, e^{(1)}, \cdots, e^{(2n-r-1)}). \tag{14}$$

If in (14) we replace e_{ik} by e_{ii} we shall have $r+1$ idempotent units participating, and hence by the assumption of the lemma

$$S(e_{11}, e_{22}, \cdots, e_{rr}, e_{ii}, e^{(1)}, \cdots, e^{(2n-r-1)}) = 0. \tag{15}$$

If further in Case II we replace e_{11} by e_{i1}, we shall have two equal units, and hence by Remark 2

$$S(e_{i1}, e_{22}, \cdots, e_{rr}, e_{i1}, e^{(1)}, \cdots, e^{(2n-r-1)}) = 0. \tag{16}$$

From (14) and (15) we obtain by addition

$$\pm S(e_{i_1 j_1}, \cdots, e_{i_{2n} j_{2n}}) = S(e_{11}, \cdots, e_{rr}, c, e^{(1)}, \cdots, e^{(2n-r-1)}), \tag{17}$$

while in Case II subtraction of (16) from (14) yields

$$\pm S(e_{i_1 j_1}, \cdots, e_{i_{2n} j_{2n}}) = S(d, e_{22}, \cdots, e_{rr}, e_{i1}, e^{(1)}, \cdots, e^{(2n-r-1)}). \tag{18}$$

In Case II our system of $2n$ units may be written in the form

$$e_{11}, \cdots, e_{rr}, e_{i1}, e^{(1)}, \cdots, e^{(2n-r-1)}. \tag{19}$$

Since we have chosen e_{i1} so that (11) holds, it follows that if in (19) we replace e_{i1} by e_{ii} and e_{11} by e_{i1}, we shall get a system of $2n$ units with an idempotent e_{ii} such that $f(i) \leqq 4$, and hence by Lemma 2

$$S(e_{i1}, e_{22}, \cdots, e_{rr}, e_{ii}, e^{(1)}, \cdots, e^{(2n-r-1)}) = 0. \tag{20}$$

From (15) and (20) we obtain by subtraction

$$S(d, e_{22}, \cdots, e_{rr}, e_{ii}, e^{(1)}, \cdots, e^{(2n-r-1)}) = 0 \tag{21}$$

while from (18) and (21) we get by addition

$$\pm S(e_{i_1 j_1}, \cdots, e_{i_{2n} j_{2n}}) = S(d, e_{22}, \cdots, e_{rr}, c, e^{(1)}, \cdots, e^{(2n-r-1)}). \tag{22}$$

Transforming now (17), respectively (22), with the matrix a defined

above, we obtain therefore in Case I as well as in Case II in view of the properties of a in relation to the three orthogonal systems (12), (15), and (14)

$$\pm a^{-1}S(e_{i_1j_1}, \cdots, e_{i_{2n}j_{2n}})\cdot a = S(e_{11}, e_{22}, \cdots, e_{rr}, e_{r+1\,r+1}, a^{-1}e^{(1)}a, \cdots, a^{-1}e^{(2n-r-1)}a). \tag{23}$$

Now express each of the matrices $a^{-1}e^{(j)}a$, $j=1, \cdots, 2n-r-1$, as a linear form of the n^2 units e_{st}, then by Remark 1 we shall be able to represent the right member of (23) as a sum of terms each having the form

$$\alpha S(e_{11}, \cdots, e_{rr}, e_{r+1\,r+1}, \cdots, e_{st}, \cdots) \tag{24}$$

where α is in the underlying field. Since the system of $2n$ units participating in (24) contains at least $r+1$ idempotents, it follows by the assumption of our lemma that (24) vanishes. Hence the left member of (23) is equal to 0, and this in turn shows that also (6) is equal to 0, q.e.d.

LEMMA 4. *For each n the standard identity $S_{2n}(x)=0$ is satisfied by each set* (5) *of $2n$ units.*

PROOF. We prove this lemma by induction. For $n=1$ this is trivial. Suppose now that $n>1$ and that the lemma is true for each positive integer less than n, then we have to show that (6) vanishes for an arbitrary system (5) of $2n$ units. Now by Remark 2 the quantity (6) vanishes if all $2n$ units in (5) are idempotent. Hence by repeatedly applying Lemma 3 it follows that (6) vanishes no matter how many of the units (5)—if any—are idempotent. This completes the proof of the lemma.

THEOREM 1 (THE MAIN THEOREM). *For each n the standard identity $S_{2n}(x)=0$ holds for the complete matric-algebra A_n.*

PROOF. Denote by $a_1, a_2, \cdots, a_{2n}$ a set of $2n$ arbitrary matrices belonging to A_n. We have to show that

$$S(a_1, a_2, \cdots, a_{2n}) = 0. \tag{25}$$

To this end express each a_i as a linear form (with coefficients in the underlying field) of the n^2 units e_{jk}. By Remark 1 we may express the left member of (25) as a sum of terms each having the form $\alpha S(e_{i_1j_1}, \cdots, e_{i_{2n}j_{2n}})$ where α is in the underlying field. By Lemma 4 each of these terms vanishes, and hence (25) holds.

3. **Uniqueness.** Each homogeneous polynomial identity of degree

m which is linear in each of its indeterminates $x_1, \cdots, x_m$ may be written in the form

$$\sum_{(i)} \alpha_{(i)} x_{i_1} x_{i_2} \cdots x_{i_m} = 0 \tag{26}$$

where $\alpha_{(i)}$ is in F and the sum on the left of (26) ranges over all permutations (i) of m letters. Some of the $\alpha_{(i)}$ may of course be equal to 0. We prove

THEOREM 2. *If the total matric-algebra A_n satisfies an identity* (26) *with $m \leqq 2n$, then $m = 2n$ and the left member of* (26) *is but for a numerical factor belonging to F the standard polynomial $S_{2n}(x)$.*

PROOF.[5] In [4] it has been shown that a minimal polynomial of A_n is at least of degree $2n$, hence we have $m = 2n$. Let us now compare two terms of the left member of (26) which but for the coefficients may be transformed in each other by transposing a pair of adjacent factors. We may assume without loss of generality that these terms are of the form

$$\alpha x_1 x_2 \cdots x_{r-1} x_r x_{r+1} \cdots x_{2n} \quad \text{and} \quad \beta x_1 x_2 \cdots x_{r-1} x_{r+1} x_r \cdots x_{2n}. \tag{27}$$

In case r is odd, that is, $r = 2i-1$, we perform the substitution

$$\begin{aligned} x_{2j} &= e_{j\,j+1}, \quad x_{2j-1} = e_{jj} && \text{for } j < i, \\ x_{2i-1} &= x_{2i} = e_{ii}, \\ x_{2j} &= e_{jj}, \quad x_{2j-1} = e_{j-1\,j} && \text{for } j > i. \end{aligned}$$

As is readily seen, under this substitution the sum of the two terms (27) yields $\alpha e_{1n} + \beta e_{1n} = (\alpha+\beta) e_{1n}$, while each other term vanishes. Hence we have $(\alpha+\beta) e_{1n} = 0$, or

$$\alpha + \beta = 0. \tag{28}$$

In case r is even, that is, $r = 2i$, we apply the substitution

$$\begin{aligned} x_{2j} &= e_{j+1\,j+1}, \quad x_{2j-1} = e_{j\,j+1} && \text{for } j < i, \\ x_{2i} &= x_{2i+1} = e_{ii}, \\ x_{2j} &= e_{j\,j+1}, \quad x_{2j-1} = e_{j-1\,j-1} && \text{for } i < j \leqq n-1, \\ x_{2n-1} &= e_{nn}, \quad x_{2n} = e_{n1}. \end{aligned}$$

Under this substitution the sum of the two terms (27) yields $\alpha e_{11} + \beta e_{11} = (\alpha+\beta) e_{11}$, while the only other summands which are possibly not

[5] This theorem was proved independently also by G. Goldring of Jerusalem University.

equal to 0 are those which correspond to the cyclic permutations of the product $e_{12}e_{22}\cdots e_{nn}e_{n1}$. Hence it follows that the left member of (26) has the form $(\alpha+\beta)e_{11}+\sum_{i=2}^{n}\alpha_i e_{ii}$. Since (26) holds, each coefficient of this sum must vanish, and this leads again to (28). Consider now an arbitrary term $\alpha_{(i)}x_{i_1}\cdots x_{i_{2n}}$ of the left member of (26). Since each permutation (i) may be derived from the identical permutation by successive transpositions of adjacent letters, it follows in view of (28) that $\alpha_{(i)}=\pm\alpha$, where the sign is positive or negative according as the permutation (i) is even or odd. Hence $\sum_{(i)}\alpha_{(i)}x_{i_1}\cdots x_{i_{2n}}=\alpha S_{2n}(x)$, q.e.d.

We now consider a polynomial $f(x_1, x_2, \cdots, x_m)$ whose degree in one of its indeterminates, say x_1, is equal to $r>1$. By applying the transformation

$$f(y+z, x_2, \cdots, x_m)-f(y, x_2, \cdots, x_m)-f(z, x_2, \cdots, x_m)=g(y, z, x_2, \cdots, x_m) \tag{29}$$

one obtains a polynomial whose general degree is not higher than that of $f(x_1, \cdots, x_m)$ and whose degree in y as well as in z is exactly $r-1$. By successive steps of this kind one obtains a polynomial whose general degree is not higher than that of $f(x_1, x_2, \cdots, x_m)$ and which is linear in each of its indeterminates (see [1, Lemma 2]). A nonlinear polynomial will be referred to in short as almost linear if each transformation of type (29) performed on this polynomial yields either zero or a polynomial linear in each of its indeterminates. One obtains easily the following lemmas.

LEMMA 5. *A nonlinear polynomial is either almost linear or it may be transformed into an almost linear polynomial by a finite number of operations of type* (29).

LEMMA 6. *A nonlinear polynomial* $f=f(x_1, x_2, \cdots, x_m)$ *is almost linear if and only if it has the form*

$$f=\sum_{i=0}^{m}f_i \tag{30}$$

where f_0 *is of degree not greater than* 1 *in each* x_i *while each* f_i, $i>0$, *is either of degree* 2 *in* x_i *and of degree not greater than* 1 *in each* x_k, $k\neq i$, *or identically equal to* 0. *Finally,* $f_i\neq 0$ *for some* $i>0$.

We shall need also the following lemma whose proof may also be omitted.

LEMMA 7. *Suppose that* $f=f_1+f_2$, *where each monomial of* f_1, *resp.* f_2,

is of degree not less than 1, *resp. of degree* 0 *in* x_1. *If identity* $f=0$ *holds in an algebra* A, *then both identities* $f_1=0$, $f_2=0$ *hold in* A.

THEOREM 3. *Let* $h(y_1, \cdots, y_k)=0$ *be a minimal identity for the total matric-algebra* A_n *such that* $k \leqq 2n$. *If either* $n>2$ *or the underlying field is not the prime field of characteristic* 2, *then* $k=2n$ *and* h *is but for a numerical factor the standard polynomial* $S_{2n}(y)$.

PROOF. We know by Theorem 1 and by [4] that the degree of h is $2n$. The case where h is linear in each y may be dispensed with[6] by Theorem 2 with the help of Lemma 7. We therefore assume that h is not linear. Since operations of type (29) transform polynomial identities into polynomial identities, it follows in view of Lemmas 5 and 6 that our identity may be transformed into an identity (1) whose corresponding polynomial f is almost linear, that is, of the form (30). By Lemma 7 we may assume that f_i, $i \geqq 0$, is of degree not less than 1 in each x (or $f_i=0$), f_0 has general degree $2n-1$ (or $f_0=0$), and f_i, $i \geqq 1$, is of general degree $2n$ in case $f_i \neq 0$. Thus it follows that we may put $m=2n-1$. We assume finally without impairing the generality of the case that $f_1 \neq 0$. Let us apply now the transformation (29). Since each f_i, $i \neq 1$, is either equal to 0 or linear in x_1 we get

$$g(y, z, x_2, \cdots, x_{2n-1}) = f_1(y+z, x_2, \cdots, x_{2n-1}) - f_1(y, x_2, \cdots, x_{2n-1}) - f_1(z, x_2, \cdots, x_{2n-1}). \tag{31}$$

The polynomial thus obtained is evidently homogeneous and linear in each of its indeterminates, and its degree is equal to $2n$. Hence by Theorem 2

$$g(y, z, x_2, \cdots, x_{2n-1}) = \alpha_1 S(y, z, x_2, \cdots, x_{2n-1}), \quad \alpha_1 \neq 0. \tag{32}$$

Our theorem will be proved if we arrive at the contradictory conclusion that $\alpha_1=0$. To this end consider an arbitrary monomial of f_1 and write it in the form $\gamma a x_1 b x_1 c$, where $\gamma \in F$ and a, b, c are products of certain x's which are not equal to x_1. Under transformation (29) this monomial yields $\gamma a(y+z)b(y+z)c - \gamma aybyc - \gamma azbzc = \gamma aybzc + \gamma azbyc$, that is, two monomials with the same coefficient γ.[7] Since on the other hand these monomials may be transformed into each other by transposing the factors y and z, it follows in view of (32) that they must have coefficients equal to $\pm\alpha_1$ with opposite signs,

[6] It follows that if h is linear in each x, then it must be homogeneous. Otherwise we would obtain by Lemma 7 an identity of degree less than $2n$, which by [4] is impossible.

[7] Apparently neither of these two monomials can be checked against any other monomial.

that is,

$$-\gamma = \gamma = \pm \alpha_1. \tag{33}$$

In case the characteristic of F is not equal to 2, it follows from (33) that $\alpha_1 = 0$, which completes the proof of the theorem in this case. In case the characteristic is equal to 2, it follows from (33) that $\gamma = \alpha_1$, that is, all coefficients of f_1 are equal. We get a similar result for each f_i, $i \geqq 1$. Denote by $p_i = p_i(x_1, \cdots, x_{2n-1})$ for $i \geqq 1$ the sum of all $(2n)!/2$ monomials which have degree equal to 2 in x_i and degree 1 in each x_k, $k \neq i$, then $f_i = \alpha_i p_i$, a relation which holds also for $f_i = 0$ (with $\alpha_i = 0$). Thus we have

$$f = f_0 + \sum_{i=1}^{2n-1} \alpha_i p_i. \tag{34}$$

Now apply the following substitution,

$$x_{2i-1} = e_{ii}, \quad i = 1, \cdots, n; \qquad x_{2i} = e_{i\,i+1}, \quad i = 1, \cdots, n-1. \tag{35}$$

The only monomial linear in each x_i which under substitution (35) yields the unit e_{1n} is $x_1 x_2 \cdots x_{2n-1}$. Each other monomial of this kind vanishes. Hence $f_0(e_{11}, e_{12}, \cdots) = \beta e_{1n}$, where $\beta \in F$. Since under (35) we get $x_{2i}^2 = 0$, $x_{2i-1}^2 = e_{ii}$, it follows easily that $p_{2i}(e_{11}, \cdots) = 0$; $p_{2i-1}(e_{11}, \cdots) = e_{1n}$. By (34) we have $0 = f(e_{11}, \cdots) = \beta e_{1n} + \sum_{i=1}^{n} \alpha_{2i-1} e_{1n}$, which implies $\beta + \sum_{i=1}^{n} \alpha_{2i-1} = 0$ or, since the characteristic is equal to 2,

$$\beta = \sum_{i=1}^{n} \alpha_{2i-1}. \tag{36}$$

Now apply the substitution

$$x_{2i-1} = e_{i\,i+1}, \ i = 1, \cdots, n-1; \quad x_{2i} = e_{ii}, \ i = 1, \cdots, n-1; \quad x_{2n-1} = e_{n1}. \tag{37}$$

The only monomial linear in each x which under (37) yields the unit e_{11} is again $x_1 x_2 \cdots x_{2n-1}$, and hence it follows that the coefficient of e_{11} in the matrix $f_0(e_{12}, e_{22}, \cdots)$ is equal to β. Since under (37) we have $x_{2i-1}^2 = 0$, $x_{2i}^2 = e_{ii}$, it follows that $e_{11} \cdot p_{2i}(e_{12}, e_{22}, \cdots) e_{11} = e_{11}$, while $e_{11} \cdot p_{2i-1}(e_{12}, e_{22}, \cdots) e_{11} = 0$. Hence the coefficient of e_{11} in the matrix $f(e_{12}, e_{22}, \cdots)$ is $\beta + \sum_{i=1}^{n} \alpha_{2i}$ and we have

$$\beta = \sum_{i=1}^{n} \alpha_{2i}. \tag{38}$$

Finally apply the substitution

$$(39)\qquad \begin{aligned} &x_i = e_{i\,i+1}, \qquad i = 1, \cdots, n-1; \qquad x_n = e_{nn}; \\ &x_i = e_{2n-i+1,2n-i}, \qquad\qquad i = n+1, \cdots, 2n-1 \end{aligned}$$

(thus for $n=3$: $x_1=e_{12}$, $x_2=e_{23}$, $x_3=e_{33}$, $x_4=e_{32}$, $x_5=e_{21}$). The monomials $x_1 x_2 \cdots x_n \cdots x_{2n-1}$ and $x_1 x_2 \cdots x_n^2 \cdots x_{2n-1}$ are the only ones participating in the f's which yield the unit e_{11}; all other monomials yield either zero or units not equal to e_{11}. Hence it follows that the coefficient of e_{11} in the matrix $f(e_{12}, \cdots, e_{21})$ is equal to $\beta+\alpha_n$, that is,

$$(40)\qquad \beta = \alpha_n.$$

If n is odd, that is, $n=2k-1$, we have by (36) and (40)

$$(41)\qquad \alpha_{2k-1} = \sum_{i=1}^{n} \alpha_{2i-1}.$$

Since the odd indices may be arbitrarily rearranged without altering the right member of (41), we obtain

$$(42)\qquad \alpha_1 = \alpha_3 = \cdots = \alpha_{n-1}.$$

If n is even, that is, $n=2k$, we have by (38) and (40)

$$(43)\qquad \alpha_{2k} = \sum_{i=1}^{n-1} \alpha_{2i}$$

which by rearranging the even indices yields

$$(44)\qquad \alpha_2 = \alpha_4 = \cdots = \alpha_{2n-2}.$$

In case $n>2$ it follows from (42) and (44) by rearranging the indices that all coefficients are equal, $\alpha_1=\alpha_2=\cdots=\alpha_{2n-1}$. By adding (36) and (38) we have $2\beta=\sum_{j=1}^{2n-1}\alpha_j=(2n-1)\alpha_1=\alpha_1$, or $\alpha_1=0$, which is the required contradiction. It remains to deal with the case where $n\leqq 2$, the field F has characteristic 2, and F contains an element δ such that $\delta\neq 0$, $\delta\neq 1$. Since all f_i, $i\neq 1$, are linear in x_1, we have in case $n=2$ the relation $f(\delta x_1, x_2, x_3)-\delta f(x_1, x_2, x_3)=$ $=\alpha_1\delta(\delta-1)p_1(x_1, x_2, x_3)$, while for $n=1$ we obtain $f(\delta x_1)-\delta f(x_1)$ $=\alpha_1\delta(\delta-1)x_1^2$. Since $\delta(\delta-1)\neq 0$, this implies the validity of the identity $\alpha_1 p_1(x_1, x_2, x_3)=0$, respectively $\alpha_1 x_1^2=0$. By substituting $x_1=e_{11}$, $x_2=e_{12}$, $x_3=e_{22}$, respectively $x_1=1$ one gets $\alpha_1 p_1(e_{11}, e_{12}, e_{22})$ $=\alpha_1 e_{11}^2 e_{12} e_{22}=\alpha_1 e_{12}=0$, respectively $\alpha_1 1=0$, that is, in both cases the desired contradiction $\alpha_1=0$. This completes the proof of the theorem.

THEOREM 4. *If either $n>2$ or F is not the prime field of characteristic 2, if further $f(x_1, x_2, \cdots, x_m)$ is a minimal polynomial of A_n, then $m\geqq 2n$ and f has the form*

$$f(x_1, \cdots, x_m) = \sum_{(i)} \alpha_{(i)} S(x_{i_1}, \cdots, x_{i_{2n}}) \tag{45}$$

where the sum ranges over all $C_{m,2n}$ combinations (i) of $2n$ letters out of m, and the $\alpha_{(i)}$ are in the underlying field.

PROOF. If $m \leqq 2n$, then by Theorem 3 we have $m = 2n$ and $f(x_1, \cdots, x_m) = \alpha S(x_1, \cdots, x_{2n})$, in which case (45) is valid. Suppose now that $m > 2n$ and consider an arbitrary monomial $\alpha_i x_{k_1} \cdots x_{k_r}$ of F, where $\alpha \neq 0$ and each x_{k_i} is some x_j. We do not assume linearity beforehand, so that some of the factors may coincide. Suppose that $x_{i_1}, x_{i_2}, \cdots, x_{i_s}$ are the distinct factors participating in this monomial, and denote by f_1, respectively f_2, the sum of all monomials which depend on those s factors only, respectively the sum of the remaining monomials, that is, $f = f_1 + f_2$. By Lemma 7 it follows that both identities $f_1 = 0, f_2 = 0$ hold in A_n. Since the degree of f is equal to $2n$, we have $s \leqq r \leqq 2n$, and hence in view of Theorem 3 it follows that $s = 2n$ and $f_1(x_{i_1}, x_{i_2}, \cdots, x_{i_s}) = \alpha_1 S(x_{i_1}, \cdots, x_{i_{2n}})$. We may repeat now the same argument with f_2, and thus proceeding, after a finite number of steps obtain the required representation (45).

It follows readily that the $C_{m,2n}$ standard polynomials defined above are linearly independent, and since each linear combination of these polynomials yields a minimal polynomial (or zero), we have the following theorem.

THEOREM 5. *If either $n > 2$ or F is not the prime field of characteristic 2, and if $x_1, x_2, \cdots, x_m$ is a set of indeterminates, then minimal polynomials of A_n depending on the x's only exist if and only if $m \geqq 2n$. The set of all such polynomials forms (if supplemented by 0) a module over whose F dimensionality is equal to $C_{m,2n}$. As a basis for this module we may choose the $C_{m,2n}$ standard polynomials defined in the previous theorem.*

The exceptional case where $n \leqq 2$ and F is the prime field of characteristic 2 is settled by the following theorem.

THEOREM 6. *The condition of Theorems 3, 4, 5 is necessary for the validity of these theorems, or in other words: The total matric-algebras A_1 and A_2 over the prime field of characteristic 2 possess nonlinear minimal identities, for example, $x + x^2 = 0$, respectively*

$$xy^2 + yxy^2 + y^2xy + y^3x + xy^2 + y^2x = 0. \tag{46}$$

PROOF. Denote by e the unit of A_2. Each element of A_2 satisfies one of the following 4 equations

$$y^2 = 0, \qquad y^2 = e, \qquad y^2 = y, \qquad y^2 = y + e.$$

By substituting each of these equations into the polynomial on the left of (46), one finds that this polynomial vanishes for each x. This completes the proof of the theorem, the case of A_1 being trivial.

REMARK. It remains to be settled by which processes—if at all—each polynomial identity of A_n may be derived from the minimal ones.

4. **Algebras.** If A is an algebra over F, and $B = A \times G$, where G is a field containing F, then evidently each identity which holds in B holds also in A. On the other hand, if an identity linear in each of its indeterminates holds in A, then this identity holds also in B (see [1, Lemma 3]). In view of Theorems 1 and 2 we have therefore the following theorem.

THEOREM 7. *If A is a simple algebra of degree n^2 over its centre F, then the degree of a minimal identity satisfied by A is exactly $2n$. A minimal identity depending on $2n$ indeterminates only and linear in each of these indeterminates $x_1, x_2, \cdots, x_{2n}$ is uniquely determined but for a numerical factor as the standard identity $S_{2n}(x) = 0$.*

Consider now a semi-simple algebra A and suppose that A is the direct sum of the ideals A' and A'', each of which satisfies an identity (1). Then identity (1) holds also in A. Indeed, let e' and e'' be the units of A', resp. A'' so that $e = e' + e''$ is the unit of A, and consider a set of m elements $a_1, a_2, \cdots, a_m$ in A. We have then $f(a_1, a_2, \cdots, a_m) = f(e'a_1, \cdots, e'a_m) + f(e''a_1, \cdots, e''a_m)$. Since (1) holds in A' as well as in A'', each of these summands vanishes, and we have in fact $f(a_1, \cdots, a_m) = 0$. This leads to the following theorem.

THEOREM 8. *Let A be a semi-simple algebra and $A', A'', \cdots, A^{(k)}$ its simple constituents. Denote by n_i^2 the order of $A^{(i)}$ over its centre. If $n^2 = \max(n_1^2, \cdots, n_k^2)$, then the degree of any minimal polynomial of A is exactly $2n$. If a minimal polynomial depends on $2n$ indeterminates only and is linear[8] in each of these indeterminates $x_1, \cdots, x_{2n}$, then this polynomial is uniquely determined but for a numerical factor as the standard polynomial $S_{2n}(x)$.*

For an arbitrary algebra of finite dimensionality we have the following theorem.

THEOREM 9. *Let r be the index of the radical N of an algebra B, and $n^2 = \max(n_1^2, \cdots, n_k^2)$ where the n_i^2 are the orders of the simple con-*

[8] As will be shown in a following communication Theorems 3–6 also retain their validity for general semi-simple algebras.

stituents of the difference algebra $A-N$. Then B satisfies the following identity of degree $2nk$.

$$[S(x_1, \cdots, x_{2n})]^k = 0.$$

PROOF. By Theorem 8 the difference algebra $A-N$ satisfies the identity $S_{2n}(x)=0$. If $b_1, \cdots, b_{2n}$ is a set of $2n$ elements belonging to B, we evidently have the relation $S(b_1, \cdots, b_{2n}) \in N$; hence, $[S(b_1, \cdots, b_{2n})]^k = 0$.

REMARK. If x_{ij}, $i=1, \cdots, 2n$; $j=1, \cdots, k$, is a set of $2nk$ indeterminates, then evidently also the identity

$$\prod_{j=1}^{k} S(x_{1j}, \cdots, x_{2n,j}) = 0$$

holds in B. This identity is homogeneous and linear in each of its indeterminates.

BIBLIOGRAPHY

1. I. Kaplansky, *Rings with a polynomial identity*, Bull. Amer. Math. Soc. vol. 54 (1948) pp. 575–580.

2. ———, *Groups with representations of bounded degree*, Canadian Journal of Mathematics vol. 1 (1949) pp. 105–112.

3. F. W. Levi, *On skew fields of a given degree*, Indian Math. Soc. (1948) pp. 85–88.

4. J. Levitzki, *A theorem on polynomial identities*, Proceedings of the American Mathematical Society vol. 1 (1950) pp. 334–341.

JERUSALEM UNIVERSITY

ON MATRICES WHOSE CHARACTERISTIC EQUATIONS ARE IDENTICAL

W. V. PARKER

In a previous paper [1][1] it was shown that if A and C are matrices such that $ACA=0$ and B is an arbitrary matrix, then AB and $A(B+C)$ have the same characteristic equation. It is the purpose of this note to prove two theorems, each of which gives the above result as a special case.

THEOREM 1. *Let A be an $n\times m$ matrix of rank $r<n$ and let C be an $m\times n$ matrix such that $ACA=kA$ (k a scalar). If B is an $m\times n$ matrix, the characteristic equation of AB is $x^{n-r}\phi(x)=0$ and the characteristic equation of $A(B+C)$ is $x^{n-r}\phi(x-k)=0$.*

Let P and Q be nonsingular matrices such that

$$PAQ=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},\quad Q^{-1}BP^{-1}=\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix},\quad \text{and } Q^{-1}CP^{-1}=\begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix}$$

where B_1 and C_1 are square matrices of order r. Then since $ACA=kA$,

$$PACAQ=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} C_1 & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} kI_r & 0 \\ 0 & 0 \end{pmatrix}$$

and hence $C_1=kI_r$.

Also

$$PACP^{-1}=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix}=\begin{pmatrix} kI_r & C_2 \\ 0 & 0 \end{pmatrix}$$

and

$$PABP^{-1}=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}=\begin{pmatrix} B_1 & B_2 \\ 0 & 0 \end{pmatrix}$$

and hence

$$P(AB+AC)P^{-1}=\begin{pmatrix} B_1+kI_r & B_2+C_2 \\ 0 & 0 \end{pmatrix}.$$

Presented to the Society, September 2, 1949; received by the editors February 16, 1949 and, in revised form, May 19, 1949.

[1] Numbers in brackets refer to the references cited at the end of the paper.

It follows that the characteristic equation of AB is $x^{n-r}\phi(x)=0$ and the characteristic equation of $A(B+C)$ is $x^{n-r}\phi(x-k)=0$ where $\phi(x)=0$ is the characteristic equation of B_1. The result of the previous paper follows if $k=0$.

In case $k\neq 0$ it may as well be assumed that $k=1$ so that $ACA=A$. Set $U=AB$, $E=AC$ and $V=U+E$. Then $E^2=E$, $EU=U$, and $EV=V$. Hence $V^nU=(U+1)^nU$ and $U^nV=(V-1)^nV$ so that $f(V)U=f(U+1)U$ and $f(U)V=f(V-1)V$ for all polynomials $f(x)$. Let $g(x)$ be the minimum function of U and $h(x)$ the minimum function of V. Then $g(V-1)V=g(U)V=0$ and $h(U+1)U=h(V)U=0$ so that $h(x)\,|\,xg(x-1)$ and $g(x)\,|\,xh(x+1)$. Hence the minimum functions of U and V satisfy one of the four relations:

(1) $h(x)=g(x-1)$; (2) $(x-1)h(x)=xg(x-1)$;

(3) $(x-1)h(x)=g(x-1)$; (4) $h(x)=xg(x-1)$.

That all four relations are actually possible may be shown by examples.

THEOREM 2. *If M and N are square matrices such that NM (or MN) $=N^2=0$, then M and $M+N$ have the same characteristic equation.*

Since $N^2=0$, N is similar to

$$\begin{pmatrix} 0_r & I_r & 0 \\ 0_r & 0_r & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where r is the rank of N. There is no loss of generality in assuming that N is in this form. Then since $NM=0$

$$M=\begin{pmatrix} M_1 & M_2 & M_3 \\ 0_r & 0_r & 0 \\ M_4 & M_5 & M_6 \end{pmatrix}$$

where M_1 and M_2 are square matrices of order r. Any principal minor of M which contains elements of M_2 also contains a row of zeros from rows $r+1$ to $2r$. Hence the characteristic equation of M is independent of the elements of M_2. But M and $M+N$ differ only in the elements of M_2 and therefore their characteristic equations are identical. If $MN=N^2=0$, then $N'M'=N'^2=0$, so that M' and $M'+N'$ have the same characteristic equation. But the characteristic equation of any matrix is the same as that of its transposed matrix and the proof of the theorem is complete.

If $ACA=0$ then $(AC)(AB)=(AC)^2=0$ and hence AB and $A(B+C)$ have the same characteristic equation.

It is well known that if A and B are square matrices, AB and BA have the same characteristic equation and if either A or B is non-singular, the two products are similar. Roth [2] has pointed out that if both A and B are singular, the products may be similar or not.

With P and Q as defined in Theorem 1

$$PABP^{-1}=\begin{pmatrix}B_1 & B_2\\ 0 & 0\end{pmatrix}=\begin{pmatrix}B_1 & 0\\ 0 & 0\end{pmatrix}+\begin{pmatrix}0 & B_2\\ 0 & 0\end{pmatrix}=M+N_1$$

where $N_1M=N_1^2=0$ and

$$Q^{-1}BAQ=\begin{pmatrix}B_1 & 0\\ B_3 & 0\end{pmatrix}=\begin{pmatrix}B_1 & 0\\ 0 & 0\end{pmatrix}+\begin{pmatrix}0 & 0\\ B_3 & 0\end{pmatrix}=M+N_2$$

where $MN_2=N_2^2=0$.

Hence

$$AB=P^{-1}MP+P^{-1}N_1P=K+S$$

where $SA=S^2=0$ and

$$BA=QMQ^{-1}+QN_2Q^{-1}=L+T$$

where $AT=T^2=0$. Since $K=P^{-1}MP=AB-S$ and $L=QMQ^{-1}=BA-T$ the following theorem is established.

THEOREM 3. *If A and B are square matrices, there exist matrices S and T such that $SA=S^2=AT=T^2=0$ and such that $AB-S$ is similar to $BA-T$.*

BIBLIOGRAPHY

1. W. V. Parker, *On the characteristic equations of certain matrices*, Bull. Amer. Math. Soc. vol. 55 (1949) pp. 115–116.

2. W. E. Roth, *A theorem on matrices*, Amer. Math. Monthly vol. 44 (1937) p. 95.

THE UNIVERSITY OF GEORGIA

SIMPLY CONNECTED HOMOGENEOUS SPACES

DEANE MONTGOMERY

Introduction. Let G be a topological group which acts transitively, as a transformation group, on a space M. If x is any point of M, let G_x be the closed subgroup made up of all elements in G which leave x fixed. There is a natural one-to-one correspondence between points of M and points of the coset space or homogeneous space G/G_x. This one-to-one correspondence is not a homeomorphism in the most general cases, but it is clearly a homeomorphism if G is a compact Lie group, and it is also a homeomorphism if G is a Lie group and M is a manifold [1].[1] In the present note the following theorem will be proved:

THEOREM A. *If G is a connected Lie group which acts transitively on a compact manifold M, and if G_x is connected, then G contains a compact subgroup which acts transitively on M.*

Because of the facts stated in the introduction this theorem could be given an equivalent formulation in the slightly different language of homogeneous spaces, but this will be omitted. When M is simply connected, G_x must be connected so that the theorem implies the following:

COROLLARY 1. *If G is a connected Lie group which acts transitively on a compact simply connected manifold M, then G contains a compact subgroup which also acts transitively on M.*

PROOF OF THE THEOREM. The proof depends on a theorem about Lie groups which was proved in essence by Cartan, Chevalley, and Malcev which will now be stated.

If G is a connected Lie group, then G is the direct product, as a space, of a compact subgroup K and a euclidean space E.

For a proof and references to original sources see [2]. It is true that maximal compact subgroups of G exist which contain any given compact subgroup, that any two such are conjugate, and that K may be chosen as one of these maximal compact groups.

In view of this theorem let the group G of Theorem A be the direct product, as a space, of K and E

$$G = K \cdot E.$$

Received by the editors May 21, 1949.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

Since G_x is connected, G_x is the direct product of a compact subgroup L and a euclidean space F

$$G_x = L \cdot F.$$

It may be assumed that $L \subset K$.

Let f_1 be the map from G/L to G/K under which each coset of L is carried to the coset of K which contains it.

$$f_1: \ G/L \to G/K.$$

This is a fiber mapping in the sense, for example, that each point of G/K is in a neighborhood whose inverse is a product of a fiber and a cross section. Each fiber is homeomorphic to K/L. But G/K is homeomorphic to E, a euclidean space of some dimension, and when a fibering has such a base space, it is a product. Therefore G/L is homeomorphic to the topological product of K/L and E, which is written as

$$G/L = K/L \times E. \tag{1}$$

Let f_2 be the map from G/L to $G/G_x = M$ under which each coset of L is carried to the coset of G_x in which it is contained.

$$f_2: \ G/L \to G/G_x.$$

This is also a fiber mapping and each of the fibers is homeomorphic to G_x/L, that is, to F, a euclidean space. But when all the fibers are homeomorphic to euclidean space it is known that a cross section A exists. The set A is a closed set in G/L, touching each fiber of f_2 precisely once, and A is homeomorphic to the base space $G/G_x = M$.

There exists a mapping P which retracts G/L onto A

$$P: \ G/L \to A,$$

that is, P is defined on all of G/L to A and each point A is fixed under P. It follows that if a cycle in A bounds in G/L, it also bounds in A. The set A is a manifold and hence contains a cycle z, at least mod 2, such that

$$\dim z = \dim A$$

and z does not bound in A. Therefore z does not bound in G/L.

From (1) it follows that the homology properties of G/L are the same as those of K/L, and therefore $\dim M = \dim A = \dim z \leqq \dim K/L$.

Since L is a maximal compact group in G_x, it follows that

$$L = K \cap G_x.$$

Then $K(x)$, the orbit of x under K, is homeomorphic to K/L so that

$$\dim M \geqq \dim K(x) = \dim K/L \geqq \dim M.$$

Since $K(x)$ is a manifold,

$$K(x) = M$$

and this completes the proof of the theorem.

COROLLARY 2. *If G is a connected Lie group which acts transitively on a compact manifold M, and if G_x has a finite number of components, then G contains a compact subgroup which also acts transitively on M.*

COROLLARY 3. *If G is a connected Lie group which acts transitively on a compact manifold M, where M has a finite fundamental group, then G has a compact subgroup which also acts transitively on M.*

Under the hypothesis of Corollaries 2 and 3, G_x has at most a finite number of components. Let G_x^* be the identity component of G_x. Since G/G_x^* is a finite covering of G/G_x and G/G_x is compact, it follows that G/G_x^* is compact. By Theorem A, G contains a compact group K such that

$$KG_x^* = G.$$

But then

$$KG_x = G,$$

which proves Corollary 2.

As a matter of fact it is not necessary to assume in Theorem A that G is connected, for if G has a countable number of components and is transitive on M, then the identity component of G is also transitive on M. The same remark applies to the corollaries.

The Klein bottle is a homogeneous space of a connected Lie group (Mostow) but not of a compact Lie group, which shows that some such extra hypothesis as the one used here is necessary.

BIBLIOGRAPHY

1. R. Arens, *Topologies for homeomorphism groups*, Amer. J. Math. vol. 68 (1946) pp. 593–610.

2. K. Iwasawa, *On some types of topological groups*, Ann. of Math. vol. 50 (1949) pp. 507–558.

3. G. D. Mostow, *Extendability of local transformation groups and groups on surfaces*, Ann. of Math. vol. 52 (1950).

INSTITUTE FOR ADVANCED STUDY

ROOTS AND FACTORS OF ORDINALS

PHILIP W. CARRUTH

1. **Introduction.** An ordinal X is termed a *root* of an ordinal A of *index* B if $X^B = A$. In §2 of this paper we employ the Cantor normal form[1] of an ordinal and make extensive use of a result of Sherman's, [3, Theorem 2.1], to determine necessary and sufficient conditions for one ordinal to be a root of another. By means of these conditions it is shown that the set of roots of an ordinal is a closed set (that is, it contains the least upper bound of each subset of it), and thus that two ordinals either have no root in common or a greatest common root. Also an immediate consequence of these conditions is the known result [1, p. 191, (13), Theorem] that an ordinal has an infinite number of roots if and only if it is indecomposable.[2] We also give necessary and sufficient conditions for an ordinal to have a greatest root less than itself.

In §3, we use results of Jacobsthal [1 and 2] to prove a theorem of which the fact that the set of left factors of an ordinal is closed (proved by other methods in [3] and [6]) and our result in §2 that the set of roots of an ordinal is closed are simple corollaries. Although the proof in this section is shorter than that in §2, the method used in §2 has the advantage of giving additional information concerning the roots of an ordinal, and it further demonstrates the usefulness of the Cantor normal form as a tool in investigations of this sort.

§4 is concerned with the decomposition of an ordinal into the product of prime factors. An ordinal is *prime* if it is not the product of two ordinals less than itself. It is known [1, p. 183, (17), Theorem] that an ordinal is prime if and only if it is one of the following: a prime integer, one greater than a transfinite indecomposable ordinal, an indecomposable power of ω. An ordinal can be written as the product of a finite number of prime factors, but this representation is not always unique. Sieczka [4] and Jacobsthal [1, p. 184, Theorem]

Presented to the Society, October 30, 1948; received by the editors October 21, 1948 and, in revised form, May 28, 1949.

[1] See [5] for information concerning the Cantor normal form and arithmetic operations with ordinals. Henceforth in this paper, whenever an ordinal is written as the sum of powers of ω, it will be assumed that the summation is the Cantor normal form. Numbers in brackets refer to the bibliography at the end of the paper.

[2] Throughout this paper, the term *indecomposable* will mean *indecomposable with respect to the operation of addition.*

showed that every ordinal has a unique prime factorization provided that certain additional conditions are satisfied by the factorization. We give some alternative additional conditions that insure the uniqueness of the prime factorization. The prime factorization satisfying these conditions has the possible advantage of using the minimum number of factors.

2. **Powers and roots of ordinals.** Throughout this section we let $A=\sum_{i=0}^{s}\omega^{\alpha(i)}a_i>0$, $\alpha(0)=\sum_{i=0}^{r}\omega^{\gamma(i)}c_i$ or 0, $X=\sum_{i=0}^{m}\omega^{\beta(i)}b_i>1$, and $B=\sigma+n\geqq 1$, where σ is a limit ordinal or 0 and n is a non-negative integer.

We shall define three types of ordinals A. The definitions will be so made that an ordinal will have a root less than itself if and only if it is of at least one of the three types. The division of such ordinals into types is strictly formal and is only necessary so that the statements of theorems concerning conditions for one ordinal to be a root of another will not become overly involved.

A is of *type* I if:

1. $s=0$.

2. a_0 is an integral power, greater than the first, of an integer greater than 1, and/or $\alpha(0)>0$.

A is of *type* II if $s<0$ and if there exist an integer j, $0\leqq j\leqq r$, and a factor p_j of c_j such that:[3]

1. $j>0$ and/or $p_j>c_j$.

2. $(\omega^{\gamma(j)}p_j+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i)(\sum_{i=0}^{j-1}\omega^{-\gamma(j)+\gamma(i)}c_i+c_j/p_j-1)<\alpha(s)$.

It is noted that the integer j described for type II is not unique; namely, let $A=\omega^{\omega^2+2}+\omega^{\omega^2+1}$. Then j could be either 0 (in which case p_j could be taken to be 1) or 1 (in which case p_j could be taken to be 2).

A is of *type* III if $s>0$, and if there exist integers j, $0\leqq j\leqq r$, and m and n such that:

1. $j>0$ and/or $m<s$.

2. $mn=s$.

3. $\alpha(km)=\sum_{i=0}^{j-1}\omega^{\gamma(i)}c_i+\omega^{\gamma(j)}c_{j,k}+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$ for $0\leqq k\leqq n-1$ $(c_{j,0}=c_j)$.

4. $c_{j,k}/(n-k)=p_j$ is an integer for $0\leqq k\leqq n-1$.

5. $\alpha(s)=\sum_{i=0}^{j-1}\omega^{\gamma(i)}c_i$.

6. If $m>1$, $-\alpha(km)+\alpha([k-1]m+i)=-\alpha(hm)+\alpha([h-1]m+i)$ for $1\leqq k,\ h\leqq n$ and $1\leqq i\leqq m-1$.

[3] Throughout this paper, whenever the upper limit of summation for the Cantor normal form of an ordinal is less than the lower limit, it is to be understood that the ordinal is 0.

7. If $n>1$, $a_0 a_s = a_{ks}$ for $1 \leqq k \leqq n-1$.
8. If $m>1$, $a_{km+i} = a_{km+i}$ for $0 \leqq k,\ h \leqq n-1$ and $1 \leqq i \leqq m-1$.

$$A = \omega^{\omega^3+\omega^2}2 + \omega^{\omega^3+\omega+2}3 + \omega^{\omega^3+\omega}10 + \omega^{\omega^2+2}3 + \omega^{\omega^2}5$$

is an example of an ordinal of type III. In this case $j=1$, $m=2$, and $n=2$.

Condition (5) for type III implies that the integer j described for that type is unique. However, m and n are not uniquely determined; namely let $A=\omega^4+\omega^3+\omega^2+\omega+1$. Then $j=0$ and both m and n could be 2 or m could be 1 and n could be 4.

It is noted that A could be of both types II and III; namely, let $A=\omega^{\omega^3+\omega+1}+\omega^{\omega^3+\omega}$. On the other hand, $A=\omega^{\omega+1}+\omega^\omega$ is of type III but not of type II; $A=\omega^4+\omega^3$ is of type II but not of type III.

THEOREM 1. *Let B be greater than* 1, *and let $X^B=A$. Then A is of at least one of the types* I, II, *and* III. *Specifically*:

1. *If $m=0$ and/or $B=\sigma$, then A is of type* I.[4]
2. *If $m>0$, $n>0$, and $\beta(m)>0$, then A is of type* II.
3. *If $m>0$, $n>0$, and $\beta(m)=0$, then A is of type* III.

PROOF OF (1). (1) is easily verified if X and B are integers. If X is an integer and $B \geqq \omega$, then $X^{\sigma+n}=(X^\omega)^\mu X^n=\omega^\mu X^n$, where $\omega\mu=\sigma$. (It is easily verified that ω is a left factor of every limit ordinal.)

If $X \geqq \omega$, $m=0$, and $n>0$, then $(\omega^{\beta(0)}b_0)^B=\omega^{\beta(0)B}b_0$ is of type I.

If $X \geqq \omega$ and $B=\sigma$, then $(\sum_{i=0}^{m}\omega^{\beta(i)}b_i)^B=\omega^{\beta(0)\sigma}$ is of type I.

PROOF OF (2). In the proof of (2) and in further proofs we shall make use of [3, Theorem 2.1] which states that a necessary and sufficient condition that an ordinal Y be a left factor of an ordinal $D=\sum_{i=0}^{k}\omega^{\beta(i)}d_i$ is that either $0<Y<\omega^{\beta(k)}$ or $Y=\omega^{\beta(j)}e_j+\sum_{i=j+1}^{k}\omega^{\beta(i)}d_i$, where e_j is a factor of d_j and $0 \leqq j \leqq k$.

$$X^B = \omega^{\beta(0)B}b_0 + \omega^{\beta(0)(B-1)+\beta(1)}b_1 + \cdots + \omega^{\beta(0)(B-1)+\beta(m)}b_m$$

$$= \sum_{i=0}^{s}\omega^{\alpha(i)}a_i.$$

Hence $\alpha(0)=\beta(0)B$. Since $n \neq 0$, it follows from [3, Theorem 2.1] that there exist an integer j, $0 \leqq j \leqq r$, and a factor p_j of c_j such that $\beta(0)=\omega^{\gamma(j)}p_j+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$. The proof of [3, Theorem 2.1] implies that $B=\sum_{i=0}^{j-1}\omega^{-\gamma(j)+\gamma(i)}c_i+c_j/p_j$. Since $B>1$, $j>0$ and/or $p_j<c_j$. The fact that $\alpha(s)=\beta(0)(B-1)+\beta(s)$ implies that $\alpha(s)>\beta(0)(B-1)$. Therefore condition (2) for type II is satisfied.

PROOF OF (3).

[4] See also [1, p. 190, Corollary 1].

$$
\begin{aligned}
X^B = {} & \omega^{\beta(0)B}b_0 + \omega^{\beta(0)(B-1)+\beta(1)}b_1 + \cdots \\
& + \omega^{\beta(0)(B-1)+\beta(m-1)}b_{m-1} \\
& + \omega^{\beta(0)(B-1)}b_0b_m + \omega^{\beta(0)(B-2)+\beta(1)}b_1 + \cdots \\
& + \omega^{\beta(0)(B-2)+\beta(m-1)}b_{m-1} + \cdots \\
& + \omega^{\beta(0)(\sigma+1)}b_0b_m + \omega^{\beta(0)\sigma+\beta(1)}b_1 + \cdots \\
& + \omega^{\beta(0)\sigma+\beta(m-1)}b_{m-1} + \omega^{\beta(0)\sigma}b_m \\
= {} & \sum_{i=0}^{s} \omega^{\alpha(i)}a_i.
\end{aligned}
$$

It is apparent that $mn=s$. Since $\beta(0)B=\alpha(0)$ and $n\neq 0$, [3, Theorem 2.1] implies that there are an integer j, $0\leqq j\leqq r$, and a factor p_j of c_j such that $\beta(0)=\omega^{\gamma(j)}p_j+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$. It can be verified now that $B=\sum_{i=0}^{j-1}\omega^{-\gamma(j)+\gamma(i)}c_i+n$. Since $B>1$, $j>0$ and/or $m<s$.

$\beta(0)(B-k)=\alpha(km)$ for $0\leqq k\leqq n$. Therefore $\alpha(km)=\sum_{i=0}^{j-1}\omega^{\gamma(i)}c_i+\omega^{\gamma(j)}c_{j,k}+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$ and $c_{j,k}=p_j(n-k)$ for $0\leqq k\leqq n-1$.

$\beta(0)\sigma=\alpha(s)=\sum_{i=0}^{j-1}\omega^{\gamma(i)}c_i$, as can be verified by direct computation.

Since $\alpha([k-1]m+i)=\alpha(km)+\beta(i)$ for $1\leqq k\leqq n$ and $1\leqq i\leqq m-1$, if $m>1$, it is seen that condition (6) for type III is satisfied. Conditions (7) and (8) are easily verified to be satisfied.

THEOREM 2. *Let A be of type* I. *Then X is a root of A if and only if one of the following conditions is satisfied*:

1. $X^n=a_0$, *where n is a non-negative integer, and* $X<\omega$.
2. $0<\beta(0)<\omega^{\gamma(r)}$, *if* $a_0=1$.
3. $\beta(0)>0$, $m=0$, $b_0=a_0$, $\beta(0)=\omega^{\gamma(j)}p_j+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$, *where p_j is a factor of c_j and* $0\leqq j\leqq r$.

If (1) *is satisfied, the index* $B=\sum_{i=0}^{r}\omega^{1+\gamma(i)}c_i+n$. *If* (2) *is satisfied,* $B=\sum_{i=0}^{r}\omega^{-\tau+\gamma(i)}c_i$, *where τ is the degree of* $\beta(0)$ (*that is, the greatest ordinal that appears as an exponent of ω in the Cantor normal form of* $\beta(0)$). *If* (3) *is satisfied,* $B=\sum_{i=0}^{j-1}\omega^{-\gamma(j)+\gamma(i)}c_i+c_j/p_j$.

PROOF. Assume that (1), (2), or (3) is satisfied. Then using the corresponding value given for B, one can verify by direct computation that $X^B=A$.

Now assume that $X^B=A$ and that $B\neq 1$. (If $B=1$, (1) or (3) is seen to hold.) The proof of (1) of Theorem 1 implies that if X is an integer, (1) must be satisfied.

If $\beta(0)>0$, then the proof of (1) of Theorem 1 implies that either $a_0=1$ and $\beta(0)$ is a left factor of $\alpha(0)$ leaving a quotient that is a limit ordinal or that $m=0$, $a_0=b_0$, and $\beta(0)$ is a left factor of $\alpha(0)$

leaving a quotient that is not a limit ordinal. In the first case, [3, Theorem 2.1] implies that $0<\beta(0)<\omega^{\gamma(r)}$. In the second case, this theorem implies that $\beta(0)=\omega^{\gamma(j)}p_j+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$, where p_j is a factor of c_j and $0\leqq j\leqq r$.

THEOREM 3. *Let A be of type* II *but not of type* III. *Then X is a root of A if and only if $X=A$ or the following conditions are satisfied*:

1. $m=s$, $b_i=a_i$ for $0\leqq i\leqq s$.

2. $\beta(0)=\omega^{\gamma(j)}p_j+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$, *where j and p_j satisfy the requirements for the integers j and p_j described for type* II.

3. $\beta(k)=-\beta(0)(\sum_{i=0}^{j-1}\omega^{-\gamma(j)+\gamma(i)}c_i+c_j/p_j-1)+\alpha(k)$ *for* $0<k\leqq m$.

If (1), (2), *and* (3) *are satisfied, then the index* $B=\sum_{i=0}^{j-1}\omega^{-\gamma(j)+\gamma(i)}c_i+c_j/p_j$.

PROOF. Assume that (1), (2), and (3) are all satisfied. Then using the value given for B, one can verify by direct computation that $X^B=A$.

Now assume that $X^B=A$ and that $B\neq 1$. Then by Theorem 1, X is a limit ordinal and $m>0$. By the proof of (2) of Theorem 1, it is seen that (1) is satisfied and that $\beta(0)$ is a left factor of $\alpha(0)$ leaving a quotient that is not a limit ordinal. Hence [3, Theorem 2.1] implies that $\beta(0)=\omega^{\gamma(j)}p_j+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$, where p_j is a factor of c_j and $0\leqq j\leqq r$, and that B has the value as stated in our theorem. Since $B>1$, $j>0$ and/or $p_j<c_j$. The proof of (2) of Theorem 1 also implies that $\beta(0)(B-1)+\beta(i)=\alpha(i)$ for $0<i\leqq m$. Thus (2) and (3) of our theorem are seen to be satisfied.

THEOREM 4. *Let A be of type* III *but not of type* II. *Then X is a root of A if and only if $X=A$ or the following conditions are satisfied*:

1. $s=mn$, *where m satisfies the requirements for the integer m described for type* III.

2. $\beta(0)=\omega^{\gamma(j)}p_j+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$, *where j is the unique integer satisfying the requirements for the integer j described for type* III.

3. $p_j=c_{j,k}/(n-k)$ for $0\leqq k\leqq n-1$.

4. If $m>1$, $\beta(i)=-\alpha(km)+\alpha([k-1]m+i)$ for $1\leqq k\leqq n$ and $1\leqq i\leqq m-1$.

5. $\beta(m)=0$.

6. $b_0=a_0$, $b_m=a_s$, *and if* $n>1$, $b_0b_m=a_{km}$ *for* $1\leqq k\leqq n-1$.

7. If $m>1$, $b_i=a_{km+i}$ *for* $0\leqq k\leqq n-1$ *and* $1\leqq i\leqq m-1$.

If X *satisfies these seven conditions, the index* $B=\sum_{i=0}^{j-1}\omega^{-\gamma(j)+\gamma(i)}c_i+n$.

PROOF. Assume that the conditions (1) through (7) are satisfied. Using the value given for B, one can compute X^B. The normal form

for X^B will be that stated in the proof of (3) of Theorem 1. (2) and (3) of this theorem together imply that $\beta(0)(B-k)=\alpha(km)$ for $0\leqq k\leqq n$. This fact together with (4) implies that if $m>1$, $\beta(0)(B-k)+\beta(i)=\alpha([k-1]m+i)$ for $1\leqq k\leqq n$ and $1\leqq i\leqq m-1$. (6) and (7) now are sufficient to show that $X^B=A$.

Now assume that $X^B=A$ and that $B\neq 1$. Then by Theorem 1, $m>0$, $\beta(m)=0$, and B is not a limit ordinal. From the proof of (3) of Theorem 1 it is seen that $s=mn$, and $\beta(0)B=\alpha(0)$. Hence [3, Theorem 2.1] implies that $\beta(0)=\omega^{\gamma(j)}p_j+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$, where p_j is a factor of c_j and $0\leqq j\leqq r$, and that $B=\sum_{i=0}^{j-1}\omega^{-\gamma(j)+\gamma(i)}c_i+n$. Since $B>1$, $j>0$, and/or $m<s$. Since $\beta(0)(B-k)=\alpha(km)$ for $0\leqq k\leqq n-1$, by direct computation it is observed that $\alpha(km)=\sum_{i=0}^{j-1}\omega^{\gamma(i)}c_i+\omega^{\gamma(j)}c_{j,k}+\sum_{i=j+1}^{r}\omega^{\gamma(i)}c_i$ for these values of k, where $c_{j,k}=p_j(n-k)$. Since $\beta(0)(B-n)=\alpha(s)$, direct computation shows that j is the unique integer described for type III.

The proof of (3) of Theorem 1 implies that if $m>1$, $\beta(i)=-\alpha(km)+\alpha([k-1]m+i)$ for $1\leqq k\leqq n$ and $1\leqq i\leqq m-1$, and also that (6) and (7) must hold. These facts together imply that m satisfies the requirements of the integer m described for type III.

THEOREM 5. *Let A be of both types* II *and* III. *Then X is a root of A if and only if one of the following three conditions is satisfied*:

1. $X=A$.
2. *X satisfies conditions* (1), (2), *and* (3) *of Theorem* 3.
3. *X satisfies conditions* (1) *through* (7) *of Theorem* 4.

THEOREM 6. *An ordinal has a root less than itself if and only if it is of at least one of the types* I, II, *and* III.

PROOF. This is an immediate consequence of Theorems 1 through 5.

THEOREM 7. *The set of roots of an ordinal is closed.*

PROOF. Let S be the set of roots of A. If A is of type I and $\alpha(0)=0$, then S consists of a finite number of integers.

If A is of type I, $\alpha(0)>0$, $a_0>1$, then S consists of a finite number of integers and of a finite number of ordinals that are greater than or equal to ω, since in $\sum_{i=0}^{r}\omega^{\gamma(i)}c_i$ there is only a finite number of coefficients, and each coefficient has only a finite set of factors.

If A is of type I, $\alpha(0)>0$, and $a_0=1$, then S consists of all ordinals greater than 1 and less than or equal to ω raised to the power $\omega^{\gamma(r)}$ and of at most a finite number of ordinals greater than ω raised to the power $\omega^{\gamma(r)}$.

If A is of type II or of type III, S consists of a finite number of

ordinals greater than ω.

COROLLARY 1. *The set of roots of an ordinal is infinite if and only if the ordinal is indecomposable.*[5]

PROOF. This is an immediate consequence of the proof of Theorem 7 and the fact that the set of indecomposable ordinals is just the set of powers of ω.

COROLLARY 2. *A has a greatest root less than itself if and only if it is of types* I, II, *or* III *and is not equal to an indecomposable power of* ω.

PROOF. This follows immediately from Theorem 6 and the proof of Theorem 7.

THEOREM 8. *Two ordinals have no common root or a greatest common root.*

PROOF. This follows directly from Theorem 7, since the intersection (if not void) of the two sets of roots of two given ordinals will contain its least upper bound, which will be therefore the greatest common root of the two ordinals.

3. **Ordinal functions.** We begin this section by giving a few definitions and results from [1] and [2]. Following Jacobsthal [1], we let β_0 be a fixed ordinal and we let correspond to each ordinal $\alpha \geqq \beta_0$ an ordinal $w(\alpha)$ having the property that if $\alpha > \alpha' \geqq \beta_0$, then $w(\alpha) > w(\alpha')$. Now let ξ_0 and η_0 be two fixed ordinals. We let correspond to each pair of ordinals $\xi \geqq \xi_0$, $\eta \geqq \eta_0$ an ordinal $g(\xi, \eta)$ having the property that $g(\xi, \eta) > \xi$. Jacobsthal proved [1, Theorem II] that corresponding to given functions $w(\alpha)$ and $g(\xi, \eta)$ there is a fixed ordinal λ and a function $f(\alpha, \beta)$ defined for $\alpha \geqq \lambda$, $\beta \geqq 1$ with the properties:

1. $f(\alpha, 1) = w(\alpha)$.
2. $f(\alpha, \beta+1) = g(f(\alpha, \beta), \alpha)$.
3. $f(\alpha, \bar{\beta}) = \lim_{\beta} f(\alpha, \beta)$, if $\bar{\beta}$ is a limit ordinal and β on the right side of equation (3) runs through all ordinals $\beta < \bar{\beta} = \lim \beta$.

The function g is called the *stem function* of f. A function defined as f is called a *C-function*. Throughout this section, $w(\alpha)$ will denote a function defined as is the function $w(\alpha)$ of [1, Theorem II], f will denote the corresponding C-function, and g will denote the stem function of f. Whenever the symbol $f(\alpha, \beta)$ is used, it will be assumed that $\beta \geqq 1$ and that $\alpha \geqq \lambda$, where λ is the ordinal λ described in [1, Theorem II]. It also will be assumed that g satisfies the following

[5] See also [1, p. 191, (13), Theorem].

condition:

CONDITION A. If $\xi \geqq \xi' \geqq \xi_0$ and $\eta > \eta' \geqq \eta_0$, then $g(\xi, \eta) > g(\xi', \eta')$ and $g(\xi, \eta') \geqq g(\xi', \eta')$.

THEOREM 9. *Let $f(\alpha, 1) \equiv \alpha$, and let $g = f_1$ be a C-function satisfying the condition $f_1(f_1(\alpha, \beta), \gamma) = f_1(\alpha, f_1(\beta, \gamma))$. Then for fixed γ, the set of all solutions α of the equation $f(\alpha, \beta) = \gamma$ is either vacuous or closed.*

PROOF. Let Z be the set of all solutions α of the equation $f(\alpha, \beta) = \gamma$, and let Z_1 be any subset of Z. We assume that Z_1 is not vacuous. Let δ be the smallest ordinal that is greater than or equal to every ordinal in Z_1. Suppose that δ is not in Z.

Let σ be the smallest ordinal satisfying the equation $f(\alpha, \sigma) = \gamma$, for α in Z_1. Let τ be the smallest ordinal in Z_1 satisfying the equation $f(\tau, \sigma) = \gamma$.

[1, Theorem III] states that if $\beta > \beta'$, then $f(\alpha, \beta) > f(\alpha, \beta')$ and therefore that if $f(\alpha, \beta) > f(\alpha, \beta')$, then $\beta > \beta'$. [1, Theorem VII] states that if $\alpha \geqq \alpha'$, then $f(\alpha, \beta) \geqq f(\alpha', \beta)$, and if $f(\alpha, \beta) = f(\alpha', \beta)$ for $\alpha > \alpha'$, then β is a limit ordinal. Consequently it is seen that for $\alpha > \tau$ in Z_1, $f(\alpha, \sigma) = \gamma$. Since according to our supposition there is no greatest element in Z_1, [1, Theorem VII] also implies that σ is a limit ordinal.

Jacobsthal proved in [2, Theorem VIII] that under the hypotheses of our theorem, if σ is a fixed limit ordinal, then the set of all ordinals α satisfying the equation $f(\alpha, \sigma) = \gamma$ has a limit α_1, and the equation $f(\alpha_1, \sigma_1) = \gamma$ has a solution σ_1. According to our supposition $\alpha_1 \neq \delta$. Hence $\delta < \alpha_1$, and there exist two ordinals, α' and α'', such that $\alpha' < \delta < \alpha''$ and such that $f(\alpha', \sigma) = f(\alpha'', \sigma) = \gamma$. But then it follows from [1, Theorem VII] that $f(\delta, \sigma) = \gamma$. This gives us a contradiction to our supposition.

The following corollary has been proved elsewhere by other methods [6, Theorem] and [3, Corollary 2.4].

COROLLARY 1. *The set of left factors of an ordinal is closed.*

PROOF. Let $g(\alpha, \beta) = f_1(\alpha, \beta) = \alpha + \beta$, and let $f(\alpha, \beta) = \alpha\beta$.

COROLLARY 2. *The set of roots of an ordinal is closed.*

PROOF. Let $g(\alpha, \beta) = f_1(\alpha, \beta) = \alpha\beta$, and let $f(\alpha, \beta) = \alpha^\beta$.

We prove another result of which [6, Lemma] is a corollary.

THEOREM 10. *Let f satisfy the condition $g(f(\alpha, \beta), f(\alpha, n)) = f(\alpha, \beta + n)$ for finite n. Let σ be a fixed ordinal and τ a fixed limit ordinal. If ρ is a solution of the equation $g(\delta, \rho) = f(\sigma, \tau)$ and if $\delta < f(\sigma, \tau)$, then $\rho \geqq f(\sigma, \omega)$.*

PROOF. [1, Theorem VI] states that if β is given and if α is an ordinal such that $w(\alpha) \leqq \beta$, then there is a unique ordinal ξ such that $f(\alpha, \xi+1) > \beta \geqq f(\alpha, \xi)$. This implies that since τ is a limit ordinal, there exists an ordinal $\mu < \tau$ such that $f(\sigma, \mu) > \delta$. Hence $f(\sigma, \tau) \leqq g(f(\sigma, \mu), \rho)$ by Condition A.

Suppose that $\rho < f(\sigma, \omega)$. Then [1, Theorem VI] implies that there is an integer n such that $\rho < f(\sigma, n)$. Therefore $f(\sigma, \tau) < g(f(\sigma, \mu), f(\sigma, n)) = f(\sigma, \mu+n)$ by hypothesis and Condition A. Since $\mu+n < \tau$ [1, Theorem III] implies that this is a contradiction.

COROLLARY 1. *Let $f(\alpha, 1) \equiv \alpha$, and let g satisfy the two conditions, $g(g(\alpha, \beta), \gamma) = g(\alpha, g(\beta, \gamma))$ and $g(\alpha, \lim \beta) = \lim_\beta g(\alpha, \beta)$. Let σ be a fixed ordinal and τ a fixed limit ordinal. If ρ is a solution of the equation $g(\delta, \rho) = f(\sigma, \tau)$ and if $\delta < f(\sigma, \tau)$, then $\rho \geqq f(\sigma, \omega)$.*

PROOF. This is an immediate consequence of [2, Theorem VI], which states that under the hypotheses of the corollary, it follows that $g(f(\alpha, \beta), f(\alpha, \gamma)) = f(\alpha, \beta+\gamma)$.

COROLLARY 2. *Let ρ be a remainder greater than 0 of $\sigma\tau$, where τ is a limit ordinal. Then $\rho \geqq \sigma\omega$.*

COROLLARY 3. *Let ρ be a right factor greater than 1 of σ^τ, where τ is a limit ordinal. Then $\rho \geqq \sigma^\omega$.*

4. Prime factorization. We revert to the notation of §2. We also let $\tau(b)$ stand for the number of prime factors in the unique prime factorization of the integer $b > 1$. We let $\tau(1) = 0$.

THEOREM 11. *An ordinal can be decomposed uniquely into a product of a finite number of prime factors if and only if it is not a limit ordinal. If A is not a limit ordinal, the number of prime factors in the unique decomposition of A is $s + \sum_{i=0}^{s} \tau(a_i)$.*

PROOF. If A is a limit ordinal, then $A = \omega R = 2\omega R$, and so the decomposition can not be unique.

Now assume that $\alpha(s) = 0$ and that $s > 0$. If $a_s = 1$, [3, Theorem 2.1] and [1, p. 183, (17), Theorem] imply that the only prime left factor of A is $\omega^{\alpha(s-1)}+1$. In this case the quotient is $\sum_{i=0}^{s-2} \omega^{-\alpha(s-1)+\alpha(i)} a_i + a_{s-1}$.

If $a_s > 1$, [3, Theorem 2.1] implies that the only prime left factors of A are the prime integral factors of a_s. If p is a prime factor of a_s, then $A = p(\sum_{i=0}^{s-1} \omega^{\alpha(i)} a_i + a_s/p)$.

Since the quotient in each case again is not a limit ordinal and since prime factorization of integers is unique, a simple induction argument

shows that the prime factorization of A is unique. It is easily seen that the number of prime factors is $s+\sum_{i=0}^{s}\tau(a_i)$.

LEMMA. *Let an ordinal greater than* 1 *be indecomposable and not prime. Then it can be decomposed uniquely into the product,* P_1P_2, *of two prime factors if the following additional condition is imposed:* $QP_2<P_1P_2$ *for every prime ordinal* Q *that is less than* P_1.

PROOF. Let A be greater than 1, indecomposable, and not prime. Then the proof of [3, Theorem 2.1] implies that A can be decomposed into the product, Q_1Q_2, of two prime factors only if Q_2 is equal to ω raised to the power $\omega^{\gamma(r)}$. Let $Q_1=\omega^{\mu}$ or $\omega^{\mu}+1$. Then Q_1Q_2 is equal to ω raised to the power $\mu+\omega^{\gamma(r)}$. $Q_1Q_2=A$ if and only if $\mu+\omega^{\gamma(r)}=\sum_{i=0}^{r}\omega^{\gamma(i)}c_i$; that is, $\mu=\delta+\eta$, where $\delta=\sum_{i=0}^{r-1}\omega^{\gamma(i)}c_i+\omega^{\gamma(r)}(c_r-1)$, and $0\leqq\eta<\omega^{\gamma(r)}$. Hence our lemma is proved. In fact, P_2 is equal to ω raised to the power $\omega^{\gamma(r)}$, and $P_1=\omega^{\delta}$ if δ is indecomposable, and $P_1=\omega^{\delta}+1$ if δ is decomposable.

THEOREM 12. *A can be factored uniquely into the product,* $P_1P_2\cdots$, *of a finite number of prime ordinals if the following conditions are imposed:*

1. *The number,* $\tau(A)$, *of prime factors is the minimum into the product of which A can be decomposed.*

2. *If* P_2 *is a limit ordinal, then* $QP_2<P_1P_2$ *for every prime ordinal* Q *that is less than* P_1.

If A is not a limit ordinal, $\tau(A)=s+\sum_{i=0}^{s}\tau(a_i)$. *If* $\alpha(s)$ *is greater than* 0 *and indecomposable,* $\tau(A)=s+1+\sum_{i=0}^{s}\tau(a_i)$. *If* $\alpha(s)$ *is decomposable,* $\tau(A)=s+2+\sum_{i=0}^{s}\tau(a_i)$.

PROOF. If A is not a limit ordinal, the results follow from Theorem 11 and its proof.

Now let $\alpha(s)$ be greater than 0. By [3, Theorem 2.1], the only possible prime left factors of A are those prime ordinals less than or equal to $\omega^{\alpha(s)}$. If Q_1 is a prime left factor of A and if $A=Q_1B$, then $B=\sum_{i=0}^{s}\omega^{\beta(i)}a_i$, where $\beta(s)\geqq 0$. Hence in order to factor A into the product of a finite number of prime factors, A must be factored as $Q_1\cdots Q_n\sum_{i=0}^{s}\omega^{\phi(i)}a_i$, where $\phi(s)=0$ and n is a positive integer, and $Q_1\cdots Q_n=\omega^{\alpha(s)}$. To factor A into the product of the minimum number of prime factors, n must be made a minimum.

If $\alpha(s)$ is indecomposable, n will equal 1 if and only if we take Q_1 equal to $\omega^{\alpha(s)}$. Then $A=\omega^{\alpha(s)}\sum_{i=0}^{s}\omega^{-\alpha(s)+\alpha(i)}a_i=\omega^{\alpha(s)}C$. Since, by Theorem 11, C has a unique prime factorization, the factorization of A satisfying the required conditions is unique, and $\tau(A)=s+1+\sum_{i=0}^{s}\tau(a_i)$.

Now let $\alpha(s)$ be decomposable. Then $\omega^{\alpha(s)}$ is not prime, and so $n>1$. Our lemma and Theorem 11 imply that A can be factored uniquely into the product, $P_1P_2\cdots$, of $s+2+\sum_{i=0}^{s}\tau(a_i)$ prime factors, where P_2 is a limit ordinal, with the condition that QP_2 be less than P_1P_2 for every prime ordinal Q that is less than P_1.

COROLLARY. *A can be factored uniquely into the product of a finite number of prime ordinals with the single restriction that the number of prime factors be the minimum into the product of which A can be decomposed if and only if one of the following conditions is satisfied:*

1. *$\alpha(s)=0$; that is, A is not a limit ordinal.*

2. *$\alpha(s)$ is indecomposable; that is, $A=DE$, where D is an indecomposable power of ω, and E is not a limit ordinal.*

3. *$\alpha(s)$ is one greater than a decomposable ordinal; that is $A=D\omega E$, where D is a decomposable power of ω, and E is not a limit ordinal.*

PROOF. This follows from Theorem 11 and the proof of our lemma and Theorem 12. In the case where $\alpha(s)$ is decomposable, the proof of the lemma implies that $\omega^{\alpha(s)}$ can be factored uniquely into the product of two prime factors with no further restriction, if and only if $\omega^{\alpha(s)}=(\omega^{\rho}+1)\omega=\omega^{\rho+1}$, where ρ is decomposable.

BIBLIOGRAPHY

1. E. Jacobsthal, *Über den Aufbau der transfiniten Arithmetik*, Math. Ann. vol. 66 (1909) pp. 145–194.

2. ———, *Zur Arithmetik der transfiniten Zahlen*, Math. Ann. vol. 67 (1909) pp. 130–144.

3. S. Sherman, *Some new properties of transfinite ordinals*, Bull. Amer. Math. Soc. vol. 47 (1941) pp. 111–116.

4. F. Sieczka, *Sur l'unicité de la décomposition de nombres ordinaux en facteurs irreductibles*, Fund. Math. vol. 5 (1924) pp. 172–175.

5. W. Sierpiński, *Leçons sur les nombres transfinis*, Paris, 1928.

6. ———, *A property of ordinal numbers*, Bull. Calcutta Math. Soc. vol. 20 (1930) pp. 21–22.

SWARTHMORE COLLEGE

ON THE DECOMPOSITION OF ORTHOGONALITIES INTO SYMMETRIES

PETER SCHERK

1. Let $\mathfrak{F}$ be a field of characteristic $\neq 2$, and let $\mathfrak{R}_n$ denote the space of all column vectors over $\mathfrak{F}$ with n components. In the following, Greek letters denote elements of $\mathfrak{F}$, while small italics $[\neq m, n, r, s]$ stand for vectors in $\mathfrak{R}_n$, and n-rowed squared matrices over $\mathfrak{F}$ are denoted by capital letters. A prime indicates transposition.

Let G be a fixed regular symmetric matrix. Thus

$$G = G', \qquad |G| \neq 0.$$

Two vectors a and b are called perpendicular if $a'Gb = 0$. Two subspaces $\mathfrak{R}^*$ and $\mathfrak{R}^{**}$ are perpendicular if $x'Gy = 0$ for all $x \subset \mathfrak{R}^*$, $y \subset \mathfrak{R}^{**}$. Obviously, these relations are symmetric. The vectors perpendicular to a given vector respectively to a given m-space form an $(n-1)$-space, respectively $(n-m)$-space.

We call the matrix T orthogonal if it leaves the expression $x'Gy$ unchanged for all x and y. This condition is equivalent to

$$T'GT = G. \tag{1}$$

If in addition

$$\text{rank } (T - I) = 1,$$

T is called a symmetry (cf. Lemma 2; I = unit matrix).

Cartan proved that every orthogonality can be decomposed into a product of n or less than n symmetries. A proof of his theorem can be found in Dieudonné's book.[1]

The purpose of this note is to show that the minimum number of symmetries into which an orthogonality T can be decomposed is in general equal to

$$m = \text{rank } (T - I).$$

An exception occurs if and only if $G(T-I)$ is skew-symmetric. In that case, this minimum number is equal to $m+2$. For a detailed description of this case cf. the last part of this note.

2. LEMMA 1. *The following three sets of properties of a matrix A are equivalent*:

Presented to the Society, September 2, 1949; received by the editors June 14, 1949.

[1] Dieudonné, *Sur les groupes classiques*, Actualités Scientifiques et Industrielles no. 1040, Paris, 1948, pp. 20–22.

(a) *rank* $(A-I)=1$, $A^2=I$;
(b) 1 *is an* $(n-1)$*-fold*, -1 *a simple eigenvalue of* A;
(c) *There are two vectors* a *and* b *such that*

$$A = I + ab' \tag{2}$$

and

$$b'a = -2. \tag{3}$$

PROOF. Obviously

$$\begin{array}{l} \text{rank } (A - I) = 1 \\ \updownarrow \\ 1 \text{ is an } (n-1)\text{-fold eigenvalue of } A \\ \updownarrow \\ (2) \text{ holds for suitable } a \neq 0,\ b \neq 0. \end{array} \tag{4}$$

(a)→(b): The first part of (b) implies that A has exactly one other eigenvalue α, and this eigenvalue is simple. From $Ax=\alpha x$ it follows by means of (a) that

$$x = Ix = A^2x = A \cdot \alpha x = \alpha \cdot Ax = \alpha^2 x.$$

Since $x \neq 0$ and $\alpha \neq 1$, $\alpha = -1$.

(b)→(c): From our assumptions, there exists an $x \neq 0$ so that

$$0 = (A + I)x = (2I + ab')x = 2x + b'x \cdot a.$$

Thus a and x are linearly dependent. We may choose $x=a$ and obtain $(2+b'a)a=0$ and therefore (3).

(c)→(a): From (3)

$$A^2 = (I + ab')(I + ab') = I + 2ab' + a(b'a)b' = I + 2ab' - 2ab' = I.$$

We call the vector a isotropic if $a'Ga=0$.

LEMMA 2. *The following three sets of properties of a matrix* A *are equivalent*:

(a) A *is orthogonal, rank* $(A-I)=1$;
(b) *There exists a non-isotropic vector* a *such that*

$$A = I - \frac{2aa'G}{a'Ga}; \tag{5}$$

(c) A *maps some non-isotropic vector* a *on* $-a$ *and every vector perpendicular to* a *on itself.*

PROOF. We first observe that (5) is equivalent to the combined three statements (2), (3), and

(6) Ga and b are linearly dependent.

(a)→(b): By means of (4) we obtain (2). Thus, A being orthogonal, we have

$$0 = A'GA - G = (I + ba')G(I + ab') - G$$

or

$$0 = b \cdot a'G(I + ab') + Ga \cdot b'. \tag{7}$$

This formula implies (6). Substituting $Ga = \lambda b$ into (7), we obtain

$$0 = \lambda bb'(I + ab') + \lambda bb' = 2\lambda bb' + \lambda b(b'a)b'$$

hence

$$\lambda(2 + b'a) \cdot bb' = 0.$$

G being regular, $\lambda \neq 0$. Thus (3) is also satisfied.

(c)→(b): The assumptions (b) of Lemma 1 hold. This implies (2) and (3). From (2), $Ax = x$ is equivalent to $b'x = 0$, and from our assumptions, it is also equivalent to $a'Gx = 0$. Thus (6) also holds.

Obviously, (a) and (c) follow from (b).

We had defined symmetries as matrices possessing the properties (a). Thus Lemma 2 gives us two alternate definitions. From Lemma 1, we obtain the following

COROLLARY. *If A is a symmetry, then*

$$A^2 = I. \tag{8}$$

LEMMA 3.

$$\textit{rank}\,(AB - I) \leqq \textit{rank}\,(A - I) + \textit{rank}\,(B - I). \tag{9}$$

PROOF. Put

$$r = \text{rank}\,(A - I), \qquad s = \text{rank}\,(B - I).$$

The vectors x with $Ax = x$, respectively $Bx = x$, form an $(n-r)$-space, respectively $(n-s)$-space. If x lies in the intersection of these two subspaces, then $ABx = Ax = x$. Thus this intersection lies in the eigenspace of AB belonging to the eigenvalue 1. The dimension of this eigenspace is therefore greater than or equal to that of this intersection. Hence it is greater than or equal to $n - r - s$. This implies (9).

If we apply (9) repeatedly to a product T of m symmetries, we get

$$\text{rank}\,(T - I) \leqq m.$$

Thus we obtain the following

COROLLARY. *An orthogonality T cannot be the product of less than rank $(T-I)$ symmetries.*

3. LEMMA 4. *Let*

$$\text{rank } S > 1, \tag{10}$$

$$S + S' \neq 0. \tag{11}$$

Then, there exists a vector b such that

$$b'Sb \neq 0 \tag{12}$$

and

$$S + S' \neq \frac{1}{b'Sb}(Sb \cdot b'S + S'b \cdot b'S'). \tag{13}$$

PROOF. We have

$$x'(S + S')x = x'Sx + x'S'x = x'Sx + (x'Sx)' = 2x'Sx. \tag{14}$$

Thus, on account of (11), there are vectors b satisfying (12). We may assume that at least one vector b_1 exists that is a solution of both (12) and

$$S + S' = \frac{1}{b'Sb}(Sb \cdot b'S + S'b \cdot b'S'). \tag{15}$$

Put

$$c = Sb_1, \qquad d = S'b_1, \qquad \alpha = b_1'Sb_1. \tag{16}$$

Then from (12)

$$c \neq 0, \qquad d \neq 0, \qquad \alpha = b_1'c = b_1'd \neq 0, \tag{17}$$

and from (15)

$$S + S' = \frac{1}{\alpha}(cd' + dc'). \tag{18}$$

Thus, (14) and (18) imply

$$x'Sx = \frac{1}{2}x'(S + S')x = \frac{1}{\alpha}c'x \cdot d'x.$$

In particular, the quadric

$$\mathfrak{Q} \equiv x'Sx = 0$$

is identical with the pair of [not necessarily different] $(n-1)$-spaces

$c'x=0$ and $d'x=0$.

Now let b be any solution of both (12) and (15). Then Sb and $S'b$ are different from zero, and comparing (15) with (18), we see that either Sb is a multiple of c and $S'b$ is one of d, or Sb is a multiple of d, while $S'b$ is one of c.

We first consider the case that a solution b_2 of (12) and (15) exists such that Sb_2 is not a multiple of c. Then Sb_2 is a multiple of d, and the vectors c and d are linearly independent. Replacing b_2 by a suitable multiple, we may assume

$$Sb_2 = d, \qquad S'b_2 = \beta c, \qquad \beta \neq 0. \tag{19}$$

Substituting $b=b_2$ into (15), we obtain

$$S + S' = \frac{\beta}{b_2' Sb_2}(dc' + cd'),$$

and therefore, on account of (18),

$$b_2' Sb_2 = \alpha\beta.$$

Thus (19) implies

$$b_2' d = \alpha\beta, \qquad b_2' c = \alpha. \tag{20}$$

The vectors

$$S(b_1 \pm b_2) = c \pm d$$

are multiples neither of c nor of d [cf. (16) and (19)]. Hence the two vectors $b_1 \pm b_2$ cannot solve the system (12), (15). Now, from (16), (17), (19), (20)

$$\begin{aligned}(b_1 + b_2)'S(b_1 + b_2) - (b_1 - b_2)'S(b_1 - b_2) \\ = 2(b_1' Sb_2 + b_2' Sb_1) = 2(b_1' d + b_2' c) = 2(\alpha + \alpha) = 4\alpha \neq 0.\end{aligned}$$

Hence, at least one of the two vectors $b_1 \pm b_2$ satisfies (12). As it cannot satisfy (15), it is a solution of both (12) and (13).

Suppose now that every solution b of (12) and (15) is mapped by S on a multiple of c. The set of all vectors x for which Sx is a multiple of c form a subspace $\mathfrak{R}_c$ of $\mathfrak{R}_n$. From (10), $\mathfrak{R}_c$ is a proper subspace of $\mathfrak{R}_n$. Hence its dimension is not greater than $n-1$. It suffices to show that there are vectors in $\mathfrak{R}_n$ that belong neither to $\mathfrak{R}_c$ nor to the quadric $\mathfrak{Q}$.

We had $b_1 \subset \mathfrak{R}_c - \mathfrak{Q}$. Since $\mathfrak{Q}$ was a pair of $(n-1)$-spaces, there exists a vector $b_2 \subset \mathfrak{Q} - \mathfrak{R}_c$. Thus, the straight line

$$b = (1 - \lambda)b_1 + \lambda b_2 \tag{21}$$

lies neither in $\mathfrak{R}_c$ nor in $\mathfrak{Q}$. Hence it has exactly one point in common with $\mathfrak{R}_c$ and at most two points with the pair $\mathfrak{Q}$ of $(n-1)$-spaces. Altogether, the straight line (21) meets the union $\mathfrak{R}_c+\mathfrak{Q}$ in not more than three points. If $\mathfrak{F}$ is not the prime field $\mathfrak{F}_3$ of three elements, this line contains more than three points. In particular, it contains points outside of $\mathfrak{R}_c+\mathfrak{Q}$.

If $\mathfrak{F}=\mathfrak{F}_3$, then an $(n-1)$-space contains 3^{n-1} points. Since $\mathfrak{R}_c$ has a dimension not greater than $n-1$ and since $\mathfrak{R}_c$ and $\mathfrak{Q}$ have the origin in common, the set $\mathfrak{R}_c+\mathfrak{Q}$ contains less than 3.3^{n-1} points, thus fewer points than the whole $\mathfrak{R}_n$. Hence, there are points outside of $\mathfrak{R}_c+\mathfrak{Q}$.[2]

LEMMA 5. *Suppose T is orthogonal, $G(T-I)$ is not skew-symmetric, and*

$$m = \text{rank } (T - I) > 1.$$

Then there exists an orthogonality U such that

(a) *T is the product of U by a symmetry,*

(b) *rank $(U-I)=m-1$,*

(c) *$G(U-I)$ is not skew-symmetric.*

PROOF. Put

$$T_0 = T - I, \qquad S = GT_0.$$

We rewrite the orthogonality definition (1) in terms of T_0 and S:

$$T'GT - G = (T_0 + I)'G(T_0 + I) - G = 0$$

or

$$T_0'GT_0 + S + S' = 0. \tag{22}$$

G being regular, we have

$$\text{rank } S = \text{rank } T_0 = m > 1.$$

Thus S satisfies the assumptions of Lemma 4, and there is a vector b such that

$$b'Sb \neq 0 \tag{23}$$

and

[2] The proof of Lemma 4 can be simplified considerably if the case $\mathfrak{F}=\mathfrak{F}_3$ is excluded. If $\mathfrak{F}$ is the real field, this lemma is trivial. The matrix $Sb\cdot b'S+S'b\cdot b'S$, $-b'Sb(S+S')$ depends continuously on b. If it vanishes outside of the quadric $\mathfrak{Q}$ identically, then it will also vanish on $\mathfrak{Q}$. Hence $Sb\cdot b'S+S'b\cdot b'S'=0$ for all $b\subset\mathfrak{Q}$. Therefore, $\mathfrak{Q}$ would be contained in the pair of at most $(n-2)$-dimensional subspaces $Sb=0$ and $S'b=0$, which is impossible.

$$S + S' \neq \frac{1}{b'Sb}(Sb \cdot b'S + S'b \cdot b'S'). \tag{24}$$

Let $\mathfrak{R}_{n-m}$ be the eigenspace belonging to the eigenvalue 1 of T. From the regularity of G and of $T = T_0 + I$

$$x \subset \mathfrak{R}_{n-m} \leftrightarrow T_0 x = 0 \leftrightarrow Sx = 0 \leftrightarrow S'x = 0 \quad \text{[cf. (22)]}. \tag{25}$$

In particular, from (23),

$$b \not\subset \mathfrak{R}_{n-m}. \tag{26}$$

Define

$$a = T_0 b. \tag{27}$$

Then, from (25),

$$a'Gx = b'S'x = 0 \qquad \text{for all } x \subset \mathfrak{R}_{n-m}. \tag{28}$$

Furthermore, (27), (22), and (14) yield

$$a'Ga = b'T_0'GT_0b = -b'(S + S')b = -2b'Sb.$$

Thus, from (27) and (23)

$$a'Ga = -2b'Sb = -2b'Ga \neq 0. \tag{29}$$

We now put

$$A = I - \frac{2aa'G}{a'Ga}, \qquad U = AT. \tag{30}$$

Thus, A is a symmetry [cf. Lemma 2]; and with A and T, U is an orthogonality. From (8)

$$T = A^2T = AU.$$

We now verify that U also has the required properties (b) and (c). If $x \subset \mathfrak{R}_{n-m}$, then on account of (28)

$$Ux = ATx = Ax = x - \frac{2a'Gx}{a'Ga} \cdot a = x.$$

Furthermore, we obtain from (30), (27), and (29)

$$(T - A)b = \left(T_0 + \frac{2aa'G}{a'Ga}\right)b = a + \frac{2a'Gb}{a'Ga}a = a - a = 0,$$

hence

$$Ub = ATb = A^2b = b.$$

Thus the eigenspace of U belonging to the eigenvalue 1 contains both $\mathfrak{R}_{n-m}$ and b. As $b \not\subset \mathfrak{R}_{n-m}$ [cf. (26)], its dimension is not less than $n-m+1$, that is,

$$\text{rank } (U - I) \leqq m - 1.$$

Thus our assertion (b) follows by applying Lemma 3 to $T = AU$.

It remains to be shown that $G(U-I)$ is not skew-symmetric. We have

$$U - T = AT - I = \left(I - \frac{2aa'G}{a'Ga}\right)(T_0 + I) - I$$

$$= T_0 - 2\frac{T_0bb'T_0'G(T_0 + I)}{-2b'Sb} = T_0 - \frac{T_0bb'S}{b'Sb}$$

[cf. (27), (29), and (22)]. Thus

$$G(U - I) = S - \frac{Sbb'S}{b'Sb},$$

and (c) follows from (24).

THEOREM 1. *Suppose T is orthogonal and $G(T-I)$ is not skew-symmetric. Then T can be written as a product of $m = \text{rank } (T-I)$ symmetries, but not of less than m.*

PROOF. For $m=0$ and $m=1$ our statement is trivial [cf. Lemma 2]. Suppose it is proved up to $m-1 \geqq 1$. From the corollary of Lemma 3, T cannot be a product of fewer than m symmetries. Thus, our theorem follows from Lemma 5 and our induction assumption.

4. From now on we assume not only that T is orthogonal but also that $G(T-I)$ is skew-symmetric. Put $T_0 = T - I$, $m = \text{rank } T_0$. Being the rank of the skew-symmetric matrix GT_0, m is even. As the case $m=0$ is trivial, we may assume $m \geqq 2$.

From our assumptions

$$GT_0 + T_0'G = 0. \tag{31}$$

Hence

$$T'GT - G = (T_0' + I)G(T_0 + I) - G = T_0'GT_0 + GT_0 + T_0'G = 0$$

implies

$$T_0'GT_0 = 0. \tag{32}$$

We obtain from (31) and (32) $GT_0^2 + T_0'GT_0 = GT_0^2 = 0$. Since G is

regular, it follows that

$$T_0^2 = 0. \tag{33}$$

Our definition of m implies that T_0 maps the whole space $\mathfrak{R}_n$ on an m-space $\mathfrak{R}_m$, and the set of the vectors x with $T_0x=0$ forms an $(n-m)$-space $\mathfrak{R}_{n-m}$. From (33)

$$\mathfrak{R}_m \subset \mathfrak{R}_{n-m} \tag{34}$$

therefore

$$m \leqq n - m \quad \text{or} \quad m \leqq n/2.$$

If $x \subset \mathfrak{R}_m$, $y \subset \mathfrak{R}_{n-m}$, then $x = T_0z$ for some z, and $T_0y=0$. From (31)

$$x'Gy = z'T_0'Gy = -z'GT_0y = 0.$$

So, the two spaces $\mathfrak{R}_m$ and $\mathfrak{R}_{n-m}$ are perpendicular. Since the vectors perpendicular to $\mathfrak{R}_{n-m}$ form an m-space containing $\mathfrak{R}_m$, this m-space is equal to $\mathfrak{R}_m$. Hence, a vector is perpendicular to $\mathfrak{R}_{n-m}$ if and only if it lies in $\mathfrak{R}_m$. We obtain from (34) that $\mathfrak{R}_m$ is perpendicular to itself. In particular, every vector in $\mathfrak{R}_m$ is isotropic.

Let

$$A = I - \frac{2aa'G}{a'Ga} \tag{35}$$

be an arbitrary symmetry. Thus a may be any non-isotropic vector. Since it cannot lie in $\mathfrak{R}_m$, it is not perpendicular to $\mathfrak{R}_{n-m}$, and the $(n-1)$-space $\mathfrak{R}_{n-1}$ perpendicular to a does not contain $\mathfrak{R}_{n-m}$. The intersection

$$\mathfrak{R}_{n-m-1} = \mathfrak{R}_{n-1} \cdot \mathfrak{R}_{n-m} \tag{36}$$

of these two spaces is therefore an $(n-m-1)$-space.

Let $U=AT$. We first show that $\mathfrak{R}_{n-m-1}$ is the eigenspace of U belonging to the eigenvalue 1. This implies in particular that

$$\text{rank } (U - I) = m + 1. \tag{37}$$

If $x \subset \mathfrak{R}_{n-m-1}$, then from (36) and (35) $Ux = ATx = Ax = x$.

Conversely, suppose $Ux=x$. Then $Tx = A^2Tx = AUx = Ax$ and hence

$$T_0x = (T - I)x = (A - I)x = -\frac{2a'Gx}{a'Ga} \cdot a. \tag{38}$$

From (38) and (32)

$$x'T_0'GT_0x = \left(\frac{2a'Gx}{a'Ga}\right)^2 a'Ga = 0.$$

Since $a'Ga \neq 0$, this implies $a'Gx = 0$ or $x \subset \mathfrak{R}_{n-1}$. Going back to (38), we obtain $T_0x = 0$ or $x \subset \mathfrak{R}_{n-m}$. Thus $x \subset \mathfrak{R}_{n-m-1}$. This proves the above statement.

Since m is even,

$$\text{rank } G(U - I) = \text{rank } (U - I) = m + 1$$

is odd. Hence, $G(U-I)$ cannot be skew-symmetric.

From Theorem 1, U is a product of $m+1$ symmetries. Hence $T = AU$ can be written as a product of $m+2$ symmetries. Suppose we have expressed T as a product of k symmetries. Then we may put $T = AU$ where A is a symmetry (35) and U is the product of $k-1$ symmetries. Since $U = AT$, we arrive again at (37). From the corollary to Lemma 3,

$$k - 1 \geqq m + 1, \quad \text{that is,} \quad k \geqq m + 2.$$

So we have the following theorem.

THEOREM 2. *Suppose T is orthogonal and $G(T-I)$ is skew-symmetric. Let*

$$m = \textit{rank } (T - I).$$

Then

$$m \equiv 0 \pmod 2 \quad \textit{and} \quad m \leqq n/2, \tag{39}$$

and T can be decomposed into a product of $m+2$ but not fewer than $m+2$ symmetries.

In order to find these transformations $T \neq I$, we choose a basis whose $m > 0$ first vectors span $\mathfrak{R}_m$. If $m < n/2$, the next $n-2m$ vectors of this basis shall lie in $\mathfrak{R}_{n-m}$. Since

$$x'Gy = y'Gx = 0 \qquad \text{for all } x \subset \mathfrak{R}_m,\ y \subset \mathfrak{R}_{n-m},$$

G has in these coordinates the form

$$G = \begin{pmatrix} 0 & 0 & G_1' \\ 0 & G_2 & * \\ G_1 & * & * \end{pmatrix} \text{ if } m < \frac{n}{2}, \quad G = \begin{pmatrix} 0 & G_1' \\ G_1 & * \end{pmatrix} \text{ if } m = \frac{n}{2}. \tag{40}$$

Here G_1 and G_2 are regular square matrices; G_1 is m-rowed, G_2 is

$(n-2m)$-rowed and symmetric. We have

$$T_0 x \subset \mathfrak{R}_m \quad \text{for all} \quad x \subset \mathfrak{R}_n \quad \text{and} \quad T_0 y = 0 \quad \text{for all} \quad y \subset \mathfrak{R}_{n-m}.$$

Hence

$$T_0 = \begin{pmatrix} 0 & T_1 \\ 0 & 0 \end{pmatrix}$$

where T_1 is an m-rowed squared matrix. Its regularity follows from rank $T_0 = m$. We obtain

$$GT_0 = \begin{pmatrix} 0 & 0 \\ 0 & G_1 T_1 \end{pmatrix}.$$

Thus (32) and (33) are satisfied. Finally, GT_0 is skew-symmetric if and only if the same holds true of $G_1 T_1$. This leads to the following construction:

Choose $n \geqq 4$ arbitrarily, $m > 0$ according to (39) and then G according to (40). Then

$$T_0 = \begin{pmatrix} 0 & G_1^{-1} T_2 \\ 0 & 0 \end{pmatrix}$$

where T_2 may be any m-rowed regular skew-symmetric matrix.

If we take, for example, $n = 4$, $m = 2$,

$$G_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad T_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

we may put

$$G = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad T_0 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$ [3]

UNIVERSITY OF SASKATCHEWAN

[3] Professor Coxeter has made the following comment on the last example: "A space with two space-like and two time-like dimensions admits a transformation leaving a whole plane invariant although it is not merely a rotation. The explanation is, of course, that this invariant plane is isotropic."

ON THE POLYNOMIAL SOLUTIONS OF THE GENERALIZED LAMÉ DIFFERENTIAL EQUATION

MORRIS MARDEN

1. **Introduction.** The generalized Lamé differential equation with which we shall be concerned is

$$P(z)w'' + Q(z)w' + R(z)w = 0, \tag{1.1}$$

where $P(z)$, $Q(z)$, and $R(z)$ are real polynomials of degrees p, at most $p-1$, and at most $p-2$ respectively. We shall assume that $P(z)$ has p distinct zeros $a_j = b_j + ic_j$ of which the first q, $0 < q \leqq p/2$, lie in the upper half-plane, the next q are respectively the conjugate imaginaries of the first q, and the remaining $p-2q$ (if any) are real. That is,

$$c_j > 0 \qquad \text{for } 1 \leqq j \leqq q,$$
$$c_j = -c_{j-q} \qquad \text{for } q < j \leqq 2q,$$

and, if $2q < p$,

$$c_j = 0 \qquad \text{for } 2q < j \leqq p.$$

We assume further that $Q(z)$ has no factor in common with $P(z)$ and that in the development

$$\frac{Q(z)}{P(z)} = \sum_{j=1}^{p} \frac{A_j e^{i\alpha_j}}{z - a_j} \tag{1.2}$$

the coefficients satisfy the inequalities

$$A_j > 0, \quad |\alpha_j| < \pi/2 \qquad \text{for } 1 \leqq j \leqq p.$$

According to Heine (see Marden [1, §9], or Marden [2]),[1] there exist some choices $V(z)$ of the polynomial $R(z)$ corresponding to each of which equation (1.1) has a polynomial solution $S(z)$. The $S(z)$ were called *Stieltjes polynomials* by Van Vleck in honor of the man who first studied the location of their zeros relative to those of $P(z)$. For a similar reason, we shall call the $V(z)$ *Van Vleck polynomials*.

In a previous paper (Marden [2]) we obtained the following result: *For each $j = 1, 2, \cdots, q$ let there be constructed an ellipse $e(a_j)$ with the points a_j and $\bar{a}_j$ as foci and with $\cos \alpha_j$ as eccentricity. Let K be the smallest convex region which encloses both all the ellipses $e(a_j)$, j*

Presented to the Society, September 7, 1948; received by the editors December 28, 1948 and, in revised form, April 26, 1949.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

$=1, 2, \cdots, q$, *and all the real points* a_j, $j=2q+1, 2q+2, \cdots, p$.[2] *Then all the zeros of every Stieltjes polynomial* $S(z)$ *and all the zeros of the corresponding Van Vleck polynomial* $V(z)$ *lie in* K.

This result is analogous to the theorem of Lucas (see Marden [1, §6]) that all the zeros of the derivative of a polynomial $f(z)$ lie in the smallest convex region containing all the zeros of $f(z)$.

In the present paper on the zeros of the $S(z)$ and $V(z)$, we propose to develop some results which will be analogous to the following theorem of Jensen (see Marden [1, §7]).

Every non-real zero of the derivative of a real polynomial $f(z)$ *lies in or on at least one of the circles* (called the Jensen circles of $f(z)$) *which has as diameter the line segment joining a pair of conjugate imaginary zeros of* $f(z)$.

We propose to develop these results by a combination of the methods introduced in Walsh [1] and Marden [2 and 3].

2. **Two lemmas.** Let us denote by n the degree of $S(z)$, by k the number of pairs of conjugate imaginary zeros of $S(z)$, by $z_j=x_j+iy_j$ $(j=1, 2, \cdots, k)$ the zeros of $S(z)$ in the upper half-plane, and, if $k<n/2$, by z_j $(j=2k+1, 2k+2, \cdots, n)$ the real zeros of $S(z)$. Thus,

$$S(z) = (z-z_1)(z-\bar{z}_1)\cdots(z-z_k)(z-\bar{z}_k)(z-z_{2k+1})\cdots(z-z_n).$$

Also, let us associate with each pair a_j, $\bar{a}_j$ $(j=1, 2, \cdots, q)$ an ellipse $E(j, k)$ defined as follows. Its center will be at the point γ_j on the real axis such that angle $\bar{a}_j, a_j, \gamma_j$ is equal to α_j; that is,

$$\gamma_j = b_j + c_j \tan \alpha_j. \tag{2.1}$$

Its minor axis will be parallel to the y-axis and be of length $m_j=2|a_j-\gamma_j|$, and its major axis will be of length $(k^{1/2}m_j)$.

We shall now state two lemmas concerning these ellipses.

LEMMA 1. *The envelope of all the circles whose diameters are the vertical chords of the ellipse* $E(j, k)$ *is the ellipse* $E(j, k+1)$.

As this lemma may be established by elementary calculus, we shall proceed to the next lemma.

LEMMA 2. *If a non-real zero* z_1 *of* $S(z)$ *lies exterior to all the circles* $E(j, 1)$, $j=1, 2, \cdots, q$, *then it lies inside at least one of the Jensen circles* $J(z_u)$ *of* $S(z)$, $u=2, 3, \cdots, k$.

PROOF. According to equation (9,3) in Marden [1], the zero z_1, not being a point a_j, satisfies the equation

[2] If $p \leqq 2q$, reference to real points a_j should be omitted.

$$\sum_{j=1}^{p} \frac{\alpha_j}{z_1 a_j} + \sum_{j=2}^{n} \frac{2}{z_1 - z_j} = 0;$$

that is,

$$\sum_{j=1}^{q} U_j + \sum_{j=2q+1}^{p} V_j + \sum_{j=2}^{k} X_j + \sum_{j=2k+1}^{n} Y_j + Z = 0 \tag{2.2}$$

where

$$\begin{aligned} U_j &= A_j\left[e^{i\alpha_j}(z_1 - a_j)^{-1} + e^{-i\alpha_j}(z_1 - \bar{a}_j)^{-1}\right], \\ V_j &= A_j(z_1 - a_j)^{-1}, \\ X_j &= 2\left[(z_1 - z_j)^{-1} + (z_1 - \bar{z}_j)^{-1}\right], \\ Y_j &= 2(z_1 - z_j)^{-1}, \\ Z &= 2(z_1 - \bar{z}_1)^{-1}. \end{aligned}$$

Let us compute the sign of the imaginary part of each term in equation (2.2). From the equation

$$\begin{aligned} 2i \operatorname{Im} U_j &= A_j\left\{e^{i\alpha_j}\left[(z_1 - a_j)^{-1} - (\bar{z}_1 - a_j)^{-1}\right]\right. \\ &\qquad \left. + e^{-i\alpha_j}\left[(z_1 - \bar{a}_j)^{-1} - (\bar{z}_1 - \bar{a}_j)^{-1}\right]\right\} \\ &= -2iy_1 A_j\left\{e^{i\alpha_j}\left[(z_1 - a_j)(\bar{z}_1 - a_j)\right]^{-1}\right. \\ &\qquad \left. + e^{-i\alpha_j}\left[(z_1 - \bar{a}_j)(\bar{z}_1 - \bar{a}_j)\right]^{-1}\right\}, \end{aligned}$$

it follows that

$$\begin{aligned} \operatorname{sg} \operatorname{Im} U_j &= -\operatorname{sg}\left(y_1 \operatorname{Re}\left\{e^{i\alpha_j}\left[(x_1 - b_j) + i(y_1 + c_j)\right]\left[(x_1 - b_j) - i(y_1 - c_j)\right]\right\}\right) \\ &= -\operatorname{sg}\left\{y_1\left[(x_1 - b_j)^2 + y_1^2 - c_j^2 - 2c_j(x_1 - b_j)\tan\alpha_j\right]\cos\alpha_j\right\} \\ &= -\operatorname{sg}\left\{y_1\left[(x_1 - \gamma_j)^2 + y_1^2 - c_j^2 \sec^2\alpha_j\right]\right\}. \end{aligned}$$

From the definitions of V_j and Z, it follows at once that

$$\operatorname{sg} \operatorname{Im} V_j = -\operatorname{sg} y_1 \quad \text{and} \quad \operatorname{sg} \operatorname{Im} Z = -\operatorname{sg} y_1.$$

To obtain $\operatorname{Im} X_j$, we need only in U_j replace a_j by z_j and $A_j e^{\alpha_j i}$ by unity. Similarly, to obtain $\operatorname{Im} Y_j$ we need only in V_j replace a_j by z_j. Thus,

$$\begin{aligned} \operatorname{sg} \operatorname{Im} X_j &= -\operatorname{sg}\left\{y_1\left[(x_1 - x_j)^2 + y_1^2 - y_j^2\right]\right\}, \\ \operatorname{sg} \operatorname{Im} Y_j &= -\operatorname{sg} y_1. \end{aligned}$$

If, now, we suppose the zero z_1 of $S(z)$ to lie outside both all the circles $E(j, 1)$, $j = 1, 2, \cdots, q$, and all the Jensen circles $J(z_j)$ of

$S(z)$, $j=2, 3, \cdots, k$, then the imaginary part of each term on the left side of equation (2.2) would have the same sign, namely $sg(-y_1)$. That is, the left side of equation (2.2) would not vanish. Hence, if all the U_j have the sign $sg(-y_1)$, at least one X_j must have the opposite sign. That is, if z_1 is outside all the circles $E(j, 1)$, it lies inside at least one circle $J(z_j)$.

We remark in passing that Lemma 2 may be modified as follows so as to admit the possibility of both positive and negative A_j in equation (1.2).

LEMMA 2′. *If a non-real zero z_1 of $S(z)$ lies simultaneously exterior to all the circles $E(j, 1)$ corresponding to positive A_j and interior to all the circles $E(j, 1)$ corresponding to negative A_j, then it lies inside at least one Jensen circle $J(Z_u)$ of $S(z)$, $u=2, 3, \cdots, k$.*

However, as Lemma 2′ is not suited to iteration, we shall revert to our original hypothesis on the A_j and in the following sections we shall use only Lemma 2.

3. **The zeros of the Stieltjes polynomials.** By applying Lemmas 1 and 2, we shall now construct a proof of the following theorem, which, when $\alpha_j=0$ for all j, reduces to a result in Walsh [1].

THEOREM I. *Let $S(z)$ be any Stieltjes polynomial which possesses k pairs of conjugate imaginary zeros. Then each non-real zero of $S(z)$ lies in or on at least one of the ellipses $E(j, k)$, $j=1, 2, \cdots, q$.*

Let us suppose that the non-real zero z_1 of $S(z)$ lies outside all the circles $E(j, 1)$. Then, by Lemma 2 it lies inside at least one Jensen circle of $S(z)$, say $J(z_2)$, the circle which has as diameter the line-segment joining z_2 and $\bar{z}_2$. If also z_2 lies outside all the circles $E(j, 1)$, then it lies in at least one of the remaining Jensen circles of $S(z)$, say $J(z_3)$. If we continue this process, the number of remaining Jensen circles diminishes. Eventually for some u, $u \leqq k$, though z_{u-1} lies outside all the circles $E(j, 1)$ and hence say in $J(z_u)$, the point z_u lies inside at least one circle $E(j, 1)$, say $E(1, 1)$.

Now, applying Lemma 1, we infer that z_{u-1} lies in $E(1, 2)$; consequently that z_{u-2} lies in $E(1, 3)$, and so on, and finally that z_1 lies in $E(1, u)$.

Since ellipse $E(1, u)$ for $u \leqq k$ is contained in the ellipse $E(1, k)$, we have completed the proof of Theorem 1.

Concerning the derivatives of $S(z)$, we may state the following theorem.

THEOREM II. *Let $S^{(r)}(z)$ denote the rth derivative of the Stieltjes*

polynomial defined in Theorem I. *Then every non-real zero of* $S^{(r)}(z)$ *lies in or on at least one of the ellipses* $E(j, k+r)$, $j=1, 2, \cdots, q$.

To prove Theorem II, let us denote by $z_j^{(r)}$ the zeros of $S^{(r)}(z)$. According to Jensen's Theorem, each non-real zero $z_1^{(r)}$ lies in at least one Jensen circle of $S^{(r-1)}(z)$, say $J(z_1^{(r-1)})$. The zero $z_1^{(r-1)}$ lies in at least one Jensen circle of $S^{(r-2)}(z)$, say $J(z_1^{(r-2)})$, and so on. The zero z_1 must lie in at least one Jensen circle of $S(z)$, say $J(z_1)$. Since, according to Theorem I, z_1 lies in an ellipse $E(j, k)$, say $E(1, k)$, we learn again by repeated use of Lemma 1 that $z_1^{(r)}$ lies in the ellipse $E(1, k+r)$.

Theorems I and II presuppose our knowledge of the number of pairs of conjugate imaginary zeros possessed by $S(z)$. If, however, we know only the degree n of $S(z)$, we may find the following immediate corollary of these theorems more useful.

COROLLARY 1. *Let* $S(z)$ *be a Stieltjes polynomial of degree* n. *Then every non-real zero of* $S(z)$ *lies in or on at least one of the ellipses* $E(j, n/2)$ *and every non-real zero of* $S^{(r)}(z)$ *lies in at least one of the ellipses* $E(j, (2r+n)/2)$, $j=1, 2, \cdots, q$. *If* n *is odd, these ellipses may be replaced by the smaller ellipses* $E(j, (n-1)/2)$ *and* $E(j, (2r+n-1)/2)$ *respectively.*

4. The zeros of the Van Vleck polynomials. We shall now prove the following theorem.

THEOREM III. *Let* $V(z)$ *be the Van Vleck polynomial which corresponds to the Stieltjes polynomial* $S(z)$ *of Theorem* I. *Then every non-real zero* t_1 *of* $V(z)$ *lies in or on at least one ellipse* $E(j, k+2)$, $j=1, 2, \cdots, q$.

Let us consider the following three cases.

A. Point t_1 lies in or on some circle $E(j, 1)$.

B. It is one of the zeros $z_1', z_2', \cdots, z_{n-1}'$ of $S'(z)$.

C. It is neither a point z_j', nor a point in or on a circle $E(j, 1)$.

The theorem is obviously true in case A. In case B, according to Theorem II, t_1 lies in some ellipse $E(j, k+1)$ and hence in some ellipse $E(j, k+2)$. In case C, we shall use the fact that t_1 satisfies equation (9,5) of Marden [1]; namely,

$$\sum_{j=1}^{p} \frac{A_j e^{i\alpha_j}}{t_1 - a_1} + \sum_{j=1}^{n-1} \frac{1}{t_1 - z_j'} = 0. \tag{4.1}$$

The left side of (4.1) is comprised of terms of the same form as the U_j, V_j, X_j, and Y_j of equation (2.2). As in Lemma 2, we conclude therefore that t_1 must lie in at least one of the Jensen circles of $S'(z)$,

say $J(z_1')$. By Theorem II, z_1' must lie in at least one of the ellipses $E(j, k+1)$ and hence by Lemma 1 point t_1 must lie in at least one of the ellipses $E(j, k+2)$. This completes the proof of Theorem III.

If we now employ the same reasoning as for Theorem II, we may infer the following result concerning the derivatives of $V(z)$.

THEOREM IV. *Let $V^{(r)}(z)$ denote the rth derivative of the polynomial $V(z)$ of Theorem* III. *Then every non-real zero of $V^{(r)}(z)$ lies in or on at least one of the ellipses $E(j, k+2+r)$, $j=1, 2, \cdots, q$.*

Finally, we may infer at once from Theorems III and IV the following complement to Corollary 1.

COROLLARY 2. *Let $V(z)$ be the Van Vleck polynomial corresponding to a Stieltjes polynomial of degree n. Then every non-real zero of $V(z)$ lies in or on at least one of the ellipses $E(j, (n+4)/2)$ and every non-real zero of $V^{(r)}(z)$ lies in or on at least one of the ellipses, $E(j, (2r+n+4)/2)$, $j=1, 2, \cdots, q$. If n is odd, these ellipses may be replaced by the smaller ellipses $E(j, (n+3)/2)$ and $E(j, (2r+n+3)/2)$ respectively.*

BIBLIOGRAPHY

M. MARDEN

1. *The geometry of the zeros of a polynomial in a complex variable*, Mathematical Surveys, no. 3, New York, American Mathematical Society, 1949.
2. *On Stieltjes polynomials*, Trans. Amer. Math. Soc. vol. 33 (1931) pp. 934–944.
3. *On the zeros of certain real rational and meromorphic functions*, Duke Math. J. vol. 16 (1949) pp. 91–97.

J. L. WALSH

1. *On the location of the roots of Lamé's polynomials*, Tôhoku Math. J. vol. 23 (1924) pp. 312–317.

UNIVERSITY OF WISCONSIN, MILWAUKEE

ON THE ALGEBRAIC INTEGRALS OF A SYSTEM OF DIFFERENTIAL EQUATIONS OF MECHANICS

J. A. BREVES FILHO

Introduction. If in the system of differential equations

$$\frac{dx_1}{dt} = -\frac{\lambda}{2}x_1x_3 + y_3, \qquad \frac{dx_2}{dt} = \frac{\lambda}{2}x_2x_3 - \rho y_3,$$

$$\frac{dx_3}{dt} = \frac{1}{2-\lambda}(\rho y_1 - y_2),$$

$(1; \lambda, \rho)$

$$\frac{dy_1}{dt} = \rho x_1y_2 - x_3y_1, \qquad \frac{dy_2}{dt} = x_3y_2 - \rho x_2y_3,$$

$$\frac{dy_3}{dt} = \frac{1}{2}(x_2y_1 - x_1y_2),$$

the parameters λ, ρ are replaced by α, β, the resulting system will be denoted by $(1; \alpha, \beta)$. The present paper will present a study of the algebraic integrals of the system $(1; \lambda, 1)$.

The equations $(1; \lambda, 1)$ are, excepting for a change of variables, those of the motion of a solid about a fix point, in a uniform field of force, when the ellipsoid of inertia relative to the fix point is of revolution and the baricenter of the solid belongs to the equatorial plane of the same ellipsoid.

Theorems 3 and 4 are original. The demonstration of Theorem 3, presented by R. Liouville [1],[1] accepted by P. Burgatti [2], is incomplete. In fact, it is implicitly admitted that one of the polynomials belonging to the rational integral possesses terms independent of y_1, y_2, y_3. The method of demonstration of Theorem 4 was inspired from a paper of Husson [3].

For the sake of simplicity of nomenclature x will be used instead of (x_1, x_2, x_3) and y instead of (y_1, y_2, y_3) wherever possible.

Let us, "ab initio," make the following two observations about the system $(1; \lambda, 1)$:

(a) The differential equations $(1; \lambda, 1)$ partake of the following particular property of homogeneity: they are not altered when the x variables are multiplied by a constant k, the y variables by k^2; and t, by k^{-1}.

Received by the editors June 20, 1949.

[1] Numbers in brackets refer to references cited at the end of the paper.

We shall say that a function $f(x, y)$ is homogeneous (in the sense alluded to) if $f(kx, k^\lambda y) = k^m f(x, y)$.

The scalar m will be called the degree of homogeneity of the function.

If the function $f(x, y)$ is homogeneous of the degree m, its derivative df/dt, as considered from $(1; \lambda, 1)$ as a function of x, y, will be of the degree $m+1$.

In fact

$$\frac{df(kx, k^\lambda y)}{d(k^{-1}t)} = k^{m+1}\frac{df(x, y)}{dt}.$$

(b) The system $(1; \lambda, 1)$ is not altered when a permutation of x_1, x_2 is made as well as of y_1, y_2 so long as $-t$ is substituted for $+t$, or, when $-x_2, -y_2$ are substituted for $+x_2, +y_2$ as well as $-t$ for $+t$.

General case $(1; \lambda, 1)$. For the sake of completeness the following theorem is included.

THEOREM 1. *Every algebraic integral of the system* $(1; \lambda, 1)$,

$$\phi(x, y),$$

is an algebraic combination of rational integrals.

In fact, if a function ϕ is algebraic, it will be a root of an equation

$$\phi^p + A_1\phi^{p-1} + \cdots + A_{p-1}\phi + A_p = 0, \tag{2}$$

whose coefficients A_i are rational functions of x, y. Without any restriction in generality we may suppose that this equation is irreducible, as otherwise we would take one of its irreducible components instead of (2).

Since ϕ is an integral of $(1; \lambda, 1)$, we have $d\phi/dt = 0$, and therefore, by deriving (2) we have

$$\frac{dA_1}{dt}\phi^{p-1} + \cdots + \frac{dA_{p-1}}{dt}\phi + \frac{dA_p}{dt} = 0, \tag{3}$$

the coefficients dA_i/dt being rational functions of x, y.

Equation (2) being, by hypothesis, irreducible, equation (3), of a lower degree, cannot subsist.

We must, therefore, have

$$\frac{dA_i}{dt} = 0, \qquad (i = 1, 2, \cdots, p).$$

These relations show that the functions A_i are (rational) integrals of system $(1;\lambda,1)$. Owing to (2), ϕ is an algebraic combination of these integrals.

THEOREM 2. *Every rational integral of system* $(1;\lambda,1)$,

$$\frac{R(x,y)}{S(x,y)},$$

is a combination of homogeneous and rational integrals.

Let us, initially, consider the developments

$$\begin{aligned} &R(kx,k^2y)\\ &\quad=H(x,y,k)[L_0(x,y)k^p+L_1(x,y)k^{p-1}+\cdots+L_p(x,y)]\\ &\quad=H(x,y,k)L,\\ &S(kx,k^2y)\\ &\quad=H(x,y,k)[M_0(x,y)k^q+M_1(x,y)k^{q-1}+\cdots+M_q(x,y)]\\ &\quad=H(x,y,k)M, \end{aligned} \tag{4}$$

where $L_0, L_1, \cdots, M_q$ are homogeneous polynomials and $H(x,y,k)$ is the H.C.F. of $R(kx,k^2y)$ and $S(kx,k^2y)$. According to the observation (a) of the introduction,

$$\frac{R(kx,k^2y)}{S(kx,k^2y)}=\frac{L}{M}$$

is still an integral of the system $(1;\lambda,1)$, and therefore

$$\frac{d}{dt}\left(\frac{L}{M}\right)=0,$$

that is,

$$M\frac{dL}{dt}-L\frac{dM}{dt}=0,$$

whence,

$$\begin{aligned} &k^{p+q}\left(M_0\frac{dL_0}{dt}-L_0\frac{dM_0}{dt}\right)\\ &\quad+k^{p+q-1}\left(M_0\frac{dL_1}{dt}+M_1\frac{dL_0}{dt}-L_1\frac{dM_0}{dt}-L_0\frac{dM_1}{dt}\right)+\cdots=0. \end{aligned}$$

The constant k being arbitrary, the following relations are ob-

tained,

$$M_0\frac{dL_0}{dt} - L_0\frac{dM_0}{dt} = 0,$$

$$M_0\frac{dL_1}{dt} + M_1\frac{dL_0}{dt} - L_1\frac{dM_0}{dt} - L_0\frac{dM_1}{dt} = 0, \tag{5}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots,$$

which constitute a system of $p+q+1$ homogeneous linear equations on the derivatives of the $p+q+2$ polynomials $L_0, L_1, \cdots, M_q$. The matrix of the coefficients of this system is

$$\begin{Vmatrix} M_0 & 0 \cdots 0 & -L_0 & 0 \cdots & 0 \\ M_1 & M_0 \cdots 0 & -L_1 & -L_0 \cdots & 0 \\ \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\ M_q & \cdots\cdots & \cdots & \cdots & \cdots \\ 0 & M_q \cdots\cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\ 0 & 0 \cdots M_q & 0 & 0 \cdots & -L_p \end{Vmatrix}.$$

L and M being prime, their (Sylvester) resultant is not identically null. Thus the rank of this matrix must be $p+q+1$ and the system (5) will have one degree of indetermination. Under these conditions it necessarily follows from (5) that

$$\frac{dL_i}{dt} = \alpha L_i \qquad (i = 1, 2, \cdots, p;\ j = 1, 2, \cdots, q) \tag{6}$$
$$\frac{dM_j}{dt} = \alpha M_j$$

α being a suitable factor of proportionality. In fact, system (5) is identically verified for the values (6).

From the above relations (6),

$$\frac{L_i}{L_0}, \quad \frac{M_j}{L_0}$$

are (homogeneous) integrals of system $(1; \lambda, 1)$. From (4) we conclude that R/S is a rational combination of these integrals.

THEOREM 3. *Every homogeneous and rational integral of system* $(1; \lambda, 1)$,

$$\frac{R_0(x, y)}{S_0(x, y)},$$

is the quotient of the division of homogeneous and whole integrals.

Without any loss of generality, let us suppose the polynomials R_0 and S_0 to be prime.

We have

$$\frac{d}{dt}\left(\frac{R_0}{S_0}\right) = 0,$$

that is,

$$\frac{dR_0}{dt} = \frac{1}{S_0} \times \frac{dS_0}{dt} \times R_0,$$

or, simply,

$$\frac{dR_0}{dt} = \mu R_0 \tag{7}$$

where μ is a polynomial. According to observation (a) of the introduction, μ must have degree 1 of homogeneity, and therefore

$$\mu = \mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3,$$

μ_1, μ_2, and μ_3 being constants. Thus (7) is transformed into

$$\frac{dR_0}{dt} = (\mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3) R_0. \tag{8}$$

According to observation (b) of the introduction, the polynomial R_1 which is obtained from R_0 by permutation of x_1, y_1, and x_2, y_2 satisfies the relation

$$\frac{dR_1}{dt} = -(\mu_2 x_1 + \mu_1 x_2 + \mu_3 x_3) R_1. \tag{9}$$

From (8, 9) it follows that

$$\frac{dR_0R_1}{dt} = (\mu_1 - \mu_2)(x_1 - x_2) R_0 R_1. \tag{10}$$

Based on the same observation, the polynomial R_2 which is obtained from R_0R_1 by substituting $-x_3$, $-y_3$ for x_3, y_3 satisfies the relation

$$\frac{dR_2}{dt} = -(\mu_1 - \mu_2)(x_1 - x_2)R_2. \tag{11}$$

From (10, 11) it follows that

$$\frac{dR_0R_1R_2}{dt} = 0,$$

and, consequently, $R_0R_1R_2$ is an integral of system $(1; \lambda, 1)$.

Finally, one can conclude that the rational integral R_0/S_0 may be considered as the quotient of the two homogeneous and whole integrals $R_0R_1R_2$ and $S_0R_1R_2$.

Kowalewski's case $(1; 1, 1)$ [4]. It is known that the system $(1; 1, \rho)$ admits the following four algebraic integrals

$$\rho x_1x_2 + \frac{1}{2}x_2^2 + \rho y_1 + y_2 = h_1, \quad x_1y_2 + x_2y_1 + x_2y_2 = h_2, \tag{12; 1, ρ}$$

$$y_1y_2 + \rho y_2^2 = h_3, \quad (\rho x_1^2 - 2y_1)(x_2^2 - 2y_2) = h_4.$$

THEOREM 4. *The system* $(1; 1, 1)$ *does not admit any other algebraic integral distinct from the integrals* $(12; 1, 1)$.

Let

$$f(x_1, y_1, y_2, x_2, x_3, y_3) \tag{13}$$

be an algebraic integral of system $(1; 1, 1)$. We shall prove that f is an algebraic combination of the integrals $(12; 1, 1)$.

The system $(1; 1, \rho)$ admits, whatever be the value of ρ, the algebraic integral

$$f(\rho x_1, \rho y_1, \rho y_2, x_2, x_3, y_3)$$

or, owing to $(12; 1, \rho)$,

$$F(h_1, h_2, h_3, h_4, \rho, x_3, y_3),$$

F being algebraic with respect to every one of the arguments.

To prove that (13) is an algebraic combination of the integrals $(12; 1, 1)$, it is sufficient to prove that F is independent of x_3 and y_3.

Let us, in the first place, notice that x_3 and y_3 are not integrals of the system $(1; 1, 1)$, therefore F cannot depend only upon x_3 or upon y_3. Let us then suppose that it depends upon x_3 and y_3.

One can always multiply F by a convenient power of ρ, so chosen that F, considered as a function of ρ, will not admit the point $\rho = 0$ neither as a zero nor as a pole. The algebraic function F of ρ is then

developable, in the domains of the value $\rho=0$, according to the rising positive powers of ρ or of $\rho^{1/p}$. In such a development the coefficients of the various powers of ρ are algebraic functions of $h_1, h_2, h_3, h_4, x_3, y_3$.

We may always suppose that for $\rho=0$, F will not be reduced to a single function of h_1, h_2, h_3, h_4. In fact, let us carry on the development of F according to the powers of $\rho^{1/p}$, until the first term whose coefficient be not reduced to a single function of h_1, h_2, h_3, h_4. Let it be

$$F = \phi(h_1, h_2, h_3, h_4, \rho) + \rho^{k/p}F_0(h_1, h_2, h_3, h_4, \rho) + \rho^{(k+1)/p}F_1 + \cdots .$$

Substituting for F the expression

$$\frac{F - \phi(h_1, h_2, h_3, h_4, \rho)}{\rho^{k/p}},$$

which means subtracting from the integral f an algebraic function of the classic integrals, we see that one may always suppose

$$F = F_0(h_1, h_2, h_3, h_4, x_3, y_3) + \rho^{1/p}F_1 + \cdots \tag{14}$$

where F_0 is not independent of x_3 and y_3.

The system (1; 1, 0) admits the algebraic integrals (12; 1, 0) and must also admit the following algebraic integral

$$F_0(h_1, h_2, h_3, h_4, x_3, y_3)$$

which results from (14) for $\rho=0$. Let us seek for this integral.

From (12; 1, 0) one gets

$$\begin{aligned} 2y_2 &= 2h_1 - x_3^2 \\ x_1y_2 - x_2y_1 &= h_2 - x_3y_3 - h(2h_1 - x_3^2)^{-1/2} \end{aligned} \tag{15}$$

where $h=2[(4h_3-h_4)h_3]^{1/2}$.

From the third and the last equations of (1; 1, 0) it follows that

$$\frac{dx_3}{2y_2} = \frac{dy_3}{x_1y_2 - x_2y_1}.$$

Owing to (15), this equation becomes

$$\frac{dx_3}{2h_1 - x_3^2} = \frac{dy_3}{h_2 - x_3y_3 - h(2h_1 - x_3^2)^{-1/2}}.$$

The solution obtained is

$$y_3(2h_1 - x_3)^{-1/2} + \int [h(2h_1 - x_3^2)^{-1/2} - h_2](2h_1 - x_3^2)^{-3/2}dx_3 = c,$$

which is a nonalgebraic integral of the system (1; 1, 0).

This conclusion contradicts the fact of F_0 being an algebraic function, therefore the initial hypothesis that assumes F to be dependent upon x_3 and y_3 is absurd.

F being independent of x_3 and y_3, it follows, as initially said, that any algebraic integral of the system (1; 1, 1) is an algebraic combination of the four integrals (12; 1, 1), that is, such a system does not admit any other algebraic integral distinct from the integrals (12; 1, 1).

References

1. R. Liouville, Acta Math. vol. 20 (1897) pp. 244–247.
2. P. Burgatti, Rend. Circ. Mat. Palermo vol. 29 (1910) p. 369.
3. Ed. Husson, Acta Math. vol. 31 (1907) p. 71.
4. S. Kowalewski, Acta Math. vol. 12 (1889) p. 177.

University of São Paulo

NOTE ON THE FUNCTIONAL EQUATIONS

$$f(xy) = f(x) + f(y), f(x^n) = nf(x)$$

JOSEPH MILKMAN

Introduction. M. Augustin-Louis Cauchy in his famous Cours d'Analyse de l'école Royale Polytechnique (Part I, *Analyse algébrique*, chap. 5, p. 109), published in 1821, gives a beautiful proof that the only continuous solution of the functional equation $f(x)+f(y)=f(xy)$, where $f(x)$ is defined for all real numbers x, is the function $f(x)=a \ln x$. Cauchy's proof reduces the equation to the Cauchy equation $f(x)+f(y)=f(x+y)$. In 1905 G. Hamel in the Mathematische Annalen proved that the discontinuous solutions of Cauchy's equation are totally discontinuous. In 1919 B. Blumberg proved that the only measurable function satisfying Cauchy's equation is the function Ax, hence the only measurable solution of

$$(1) \qquad f(xy) = f(x) + f(y), \qquad f(x^n) = nf(x)$$

is $a \ln x$. In the past thirty years there have been a number of papers on these functional equations where the domain of the variable has usually been the real number system. In 1946 Erdös proved that if

Presented to the Society, October 29, 1949; received by the editors May 23, 1949.

$f(m)$ is additive and $f(m+1) \geqq f(m)$ then $f(m) = c \ln m$ (Ann. of Math. vol. 47 (1946)). This implies Theorem I of this paper. Professor H. Shapiro of New York University suggested that a more elementary proof of Theorem I than Erdös's should be possible and this led me to the proof of Theorem I, given below, in which I show that any solution $f(m)$ can be imbedded in a continuous solution $f(x)$, thus rendering the theorem a corollary of Cauchy's. I wish to express my thanks to Professor Freidman of New York University for some suggestions he made regarding this paper.

THEOREM I. *If $f(m+1) > f(m)$ and $f(mn) = f(m) + f(n)$ hold for all positive integers m and n, then $f(m) = c \ln m$.*

PROOF. If m/n, m and n are positive integers, $f(m/n) = f(m) - f(n)$ since $m = (m/n) \cdot n$. If m and n are relatively prime positive integers, we define $f(m/n) = f(m) - f(n)$. If λ is an integer and m and n are relatively prime, $f(m/n) = f(m) + f(\lambda) - f(n) - f(\lambda) = f(\lambda m) - f(\lambda n)$, so that if we define $f(m/n) = f(m) - f(n)$ for any two integers m and n, $f(m/n)$ is a function of the rational number m/n.

If a, b, c, and d are positive integers such that $a/b > c/d$, then $f(a/b) > f(c/d)$, since $f(a/b) = f(ad) - f(bd)$. $f(c/d) = f(bc) - f(bd)$ and $f(ad) > f(bc)$.

If a, b, c, and d are positive integers,

$$\begin{aligned} f(ac/bd) &= f(ac) - f(bd) \\ &= f(a) + f(c) - f(b) - f(d) \\ &= f(a/b) + f(c/d). \end{aligned}$$

$f(m) = f(m) + f(1)$, therefore $f(1) = 0$.

If $b_i > 1$ and $\{b_i\}$ is a sequence having 1 as a limit where b_i is a rational number, for i sufficiently large $f(b_1) > f(b_i) > f(1) = 0$. Therefore there exists a monotonically decreasing subsequence $\{b_i'\}$ of $\{b_i\}$ having 1 as a limit and a corresponding monotonically decreasing sequence $\{f(b_i)\}$ having a real number f as a limit. Suppose $f \neq 0$. Let N be an integer such that $Nf > f(2)$; let b_λ' be a b_i' such that $1 < b_\lambda' < 2^{1/2N}$, then $1 < b_\lambda'^N < 2$,

$$f(b_\lambda'^N) = Nf(b_\lambda')$$

or

$$f(b_\lambda') = \frac{f(b_\lambda'^N)}{N} < \frac{f(2)}{N} < f$$

contrary to the choice of $f(b_i')$ as a monotonically decreasing sequence approaching f as a limit. Therefore $f=0$. Since f was any limit point of the sequence $\{f(b_i)\}$, we have the result that for any sequence of rationals $b_i>1$ such that $b_i\to 1$, $f(b_i)\to 0$. If r is rational and positive

$$f(1/r) = f(1) - f(r) = -f(r).$$

Therefore if $b_i\to 1$ and $b_i<1$, $f(b_i)\to 0$. From the above we conclude that if $b_i\to 1$, $f(b_i)\to 0$.

If $\{a_i\}$ is a convergent sequence of positive rationals $f(a_m)-f(a_n)=f(a_m/a_n)\to 0$ as $|a_m-a_n|\to 0$, for as $|a_m-a_n|\to 0$, $a_m/a_n\to 1$, therefore $\{f(a_i)\}$ converges. Let $\lim_{i\to\infty} a_i=a$. If a is rational, then $a_i/a\to 1$ and $f(a_i/a)=f(a_i)-f(a)\to 0$ and hence $f(a_i)\to f(a)$. If a is irrational, define $f(a)$ as $\lim_{i\to\infty} f(a_i)$. $f(a)$ is unique for if $a_i\to a$, $f(a_i)\to f$ and if $a_i'\to a$, $f(a_i')\to f'$, then

$$f(a_i) - f(a_i') = f\left(\frac{a_i}{a_i'}\right)\to f - f',$$

but $a_i/a_i'\to 1$, hence $f-f'=0$. We have now defined $f(x)$ for all positive real numbers x so that when $x_i\to x$, $f(x_i)\to f(x)$, that is $f(x)$ is continuous. If x and y are any real numbers, let $\{x_i\}$ and $\{y_i\}$ be sequences of rational numbers such that $x_i\to x$ and $y_i\to y$, then $f(x_iy_i)=f(x_i)+f(y_i)$ and hence $f(xy)=f(x)+f(y)$ for all positive real numbers x and y. The only continuous solution to the functional equation

$$f(x) + f(y) = f(xy)$$

is

$$f(x) = c \ln x.$$

Therefore

$$f(n) = c \ln n.$$

THEOREM II. *If $f(x)$ is a monotonically increasing function defined for all positive real numbers x, such that $f(x^n)=nf(x)$ for all positive integers n, then $f(x)=k \ln (x)$ where $k=f(e)$.*

PROOF. $f(x)=nf(x^{1/n})$ or $f(x^{1/n})=1/nf(x)$. $f(x^{m/n})=(m/n)f(x)$, or if r is a rational number, $f(x^r)=rf(x)$. In particular $f(e^r)=rf(e)=f(e)\cdot\ln(e^r)$. Let x be any positive real number and let $y=\ln x$, that is $x=e^y$. Let the sequence $\{r_i\}$ be a monotonically increasing sequence of rational numbers having y as a limit, then $e^{r_i}\to x$ and $\{e^{r_i}\}$ is monotonically increasing and $\{f(e^{r_i})\}$ is monotonically increasing and has the limit of $f(e) \ln x$. If $f(x)<f(e) \ln x$, then there is an r_n such

that $f(x) < f(e^{r_n})$ with $x > e^{r_n}$ contrary to the monotonicity of $f(x)$. If $f(x) > f(e) \ln x$, let $\{r_i\}$ be a decreasing sequence of rational numbers having y as a limit, then $\{e^{r_i}\}$ is monotonically decreasing to the limit x and $\{f(e^{r_i})\}$ is a monotonically decreasing sequence having $f(e) \ln x$ as a limit. We can now find an r_m such that $f(x) > f(e^{r_m})$ with $x < e^{r_m}$, contrary to the monotonicity of $f(x)$. Therefore $f(x) = k \ln x$.

COROLLARY. *If $f(x^p) = pf(x)$ and $f(x)$ is monotonically increasing for all real numbers $x > 0$ and all prime numbers p, then $f(x) = k \ln x$.*

COROLLARY. *If $f(x^p) = pf(x)$ and $f(x)$ is monotonically decreasing for all real numbers $x > 0$, and all prime numbers p, then $f(x) = -k \ln x$.*

COROLLARY. *If $f(r)$ is a monotonically increasing function on a set S which is everywhere dense among the positive real numbers and contains all positive integral powers of all members of the set, for example, the rational numbers, and $f(r^p) = pf(r)$ for any prime p, then $f(r) = k \ln r$.*

PROOF. $f(r^n) = nf(r)$ for all r in S and positive integers n. Let R be a real number and let $\{r_i\}$ be an increasing sequence of elements of S having R as a limit, then if $\{r_i'\}$ is another such sequence, the limit L of $\{f(r_i)\}$ = the limit L' of $\{f(r_i')\}$. For suppose $L > L'$; then there exists an r_n such that $f(r_n) > L'$, and an $r_m' > r_n$, hence $L' > f(r_m') > f(r_n)$, a contradiction. Similarly, we arrive at a contradiction if we assume $L' > L$. Therefore $L = L'$. We may thus extend the domain of $f(r)$ unambiguously to all real numbers greater than 0, by defining $f(R)$ to be the limit of $\{f(r_n)\}$ for any irrational positive number R. Under this definition monotonicity of the function holds in the new extended domain of all positive real numbers. Since R^n is the limit of $\{r_i^n\}$, r_i defined above, $f(R^n) = nf(R)$ for any real number R and the corollary follows from the theorem.

UNITED STATES NAVAL ACADEMY AND
NEW YORK UNIVERSITY

TRANSITIVE SYSTEMS OF LINEAR OPERATORS ON A BANACH SPACE[1]

BERTRAM YOOD

Let $\mathfrak{X}$ be an infinite-dimensional Banach space. Mackey[2] has shown, under more general conditions, that if $x_1, \cdots, x_n, y_1, \cdots, y_n$ are any two sets of n linearly independent elements in $\mathfrak{X}$, then there exists an isomorphism T of $\mathfrak{X}$ with itself such that $T(x_i)=y_i$, $i=1, \cdots, n$. In other words the collection of isomorphisms is *transitive for linearly independent sets of elements of* $\mathfrak{X}$. This property is shared by many other sets of linear operators, for example by the set of all operators with finite-dimensional range in $\mathfrak{X}$. These two sets of linear operators are semi-groups.[3] In this note we give a condition for the transitivity property above which is necessary for all sets of linear operators and which while not in general sufficient is so for semi-groups. In this result the operators need not be assumed to be bounded or defined everywhere in $\mathfrak{X}$. $\mathfrak{E}(\mathfrak{X})$ will be used to designate the collection of all bounded linear operators with domain $\mathfrak{X}$ and range in $\mathfrak{X}$.

THEOREM 1. *Let $\mathfrak{G}$ be a collection of linear operators with domain and range in an infinite-dimensional Banach space $\mathfrak{X}$. For $\mathfrak{G}$ to be transitive for linearly independent sets of elements of $\mathfrak{X}$ it is necessary that for each finite-dimensional subspace F of $\mathfrak{X}$ there exists a $T\in\mathfrak{G}$ which is the identity on F and a number $\epsilon>0$ such that if $U\in\mathfrak{E}(\mathfrak{X})$ takes F into F and $\|U\|<\epsilon$, then there exists an operator $V\in\mathfrak{G}$ which agrees with $T+U$ on F. If $\mathfrak{G}$ is a semi-group, this condition is sufficient.*

It is readily seen that the condition is necessary. That it is not sufficient in general can be seen by taking $\mathfrak{G}$ to be an ϵ-neighborhood of the identity in the Banach space $\mathfrak{E}(\mathfrak{X})$.

Let $\mathfrak{G}$ be a semi-group. Let F be a finite-dimensional subspace of $\mathfrak{X}$ and T_0 be its associated operator in $\mathfrak{G}$, and let $x_1, \cdots, x_n$ be linearly independent elements of F. There exist x_i^* in $\mathfrak{X}^*$ (the conjugate space of $\mathfrak{X}$) such that $x_i^*(x_j)=\delta_{ij}$, $i, j=1, \cdots, n$. For elements

Received by the editors December 28, 1948 and, in revised form, June 13, 1949.

[1] This paper is based on a portion of the author's dissertation, Yale University, 1947. The author wishes to thank the referee for his suggestions which made possible an appreciable shortening of proof of Theorem 1 and clarified the nature of that proof.

[2] G. W. Mackey, *On infinite-dimensional linear spaces*, Trans. Amer. Math. Soc. vol. 57 (1945) pp. 155–207, see Theorem II-3.

[3] By a semi-group of linear operators is meant a collection which contains TU whenever it contains T and U.

$v_i \in F$ for which $\|v_i - x_i\| < \delta = \min\{1, \epsilon\}/2n \max \|x_i^*\|$, let the transformation W be determined by

$$W(x) = \sum_{i=1}^{n} x_i^*(x)(v_i - x_i).$$

Then $\|W\| < \min\{1, \epsilon\}/2$. Thus $T_0 + W$ is an isomorphism on F. The operator $U = \sum_{i=1}^{n} (-1)^i W^i$ has the properties $\|U\| < \epsilon$ and $(T_0 + U)(T_0 + W)x = x$ for all x in F. By hypothesis there exist transformations U_1 and W_1 in $\mathfrak{G}$ which agree with $T_0 + U$ and $T_0 + W$ respectively in F. Then by the above $W_1(x_i) = v_i$ and $U_1(v_i) = x_i$, $i = 1, \cdots, n$. If also for each i, $\|v_i' - x_i\| < \delta$, $v_i' \in F$, we may define $W_1' \in \mathfrak{G}$ as above so that $W_1'(x_i) = v_i'$. Then $W_1' U_1(v_i) = v_i'$, $i = 1, \cdots, n$. The semi-group property of $\mathfrak{G}$ shows that (a) for each F and $x_1, \cdots, x_n$ linearly independent in F there exists a $\delta > 0$ such that for every v_i, w_i in F with $\|v_i - x_i\| < \delta$ and $\|w_i - x_i\| < \delta$, there exists V in $\mathfrak{G}$ with $V(v_i) = w_i$, $i = 1, \cdots, n$.

Next for a fixed F and $n \leqq \dim F$ consider the Cartesian product space $\overline{F} = F \times \cdots \times F$ (n factors) made up of elements of the form $\bar{x} = (x_1, \cdots, x_n)$ with $\|\bar{x}\| = \max_{i \leqq n} \|x_i\|$. $\overline{F}$ is a finite-dimensional Banach space.[4] Let $\overline{L}$ be the set of all $\bar{x} \in \overline{F}$ such that $x_1, \cdots, x_n$ are linearly independent in F. By (a), $\overline{L}$ is open in $\overline{F}$ and to each point $\bar{x}$ of $\overline{L}$ there exists a $\delta > 0$ such that $\|\bar{x} - \bar{v}\| < \delta$ and $\|\bar{x} - \bar{w}\| < \delta$ imply the existence of a transformation $V \in \mathfrak{G}$ such that $\overline{V}(\bar{v}) = \bar{w}$ where $\overline{V}(\bar{v}) = (V(v_1), \cdots, V(v_n))$. Let $\overline{K}$ be the set of $\bar{y}$ in $\overline{L}$ which can be reached from $\bar{x}$ along a finite chain of such spheres each overlapping its predecessor. Then by (a) for such a $\bar{y}$ there exists a sequence $V_1, \cdots, V_r$ in $\mathfrak{G}$ such that for the product $W = V_r \cdots V_1$, $\overline{W}(\bar{x}) = \bar{y}$ and by the semi-group property, $W \in \mathfrak{G}$.

From this it follows by a standard topological argument, used in connection with analytic continuation and elsewhere,[5] that the set $\overline{K}$ is open and closed in $\overline{L}$. Therefore (b) if $\bar{x}$ and $\bar{y}$ are in the same component of $\overline{L}$, there is a $V \in \mathfrak{G}$ such that $\overline{V}(\bar{x}) = \bar{y}$.

Call $\bar{x}$ and $\bar{y}$ *fully independent* if $x_1, \cdots, x_n, y_1, \cdots, y_n$ form a linearly independent set of $2n$ elements and suppose that $\bar{x}, \bar{y} \in \overline{L}$. For such a pair $\bar{x}$ and $\bar{y}$, and for any real α, we have $\alpha\bar{x} + (1-\alpha)\bar{y} \in \overline{L}$, for $\sum_{i=1}^{n} \beta_i(\alpha x_i + (1-\alpha)y_i) = 0$ implies $\beta_i = 0$, $i = 1, \cdots, n$. Thus the line segment joining $\bar{x}$ and $\bar{y}$ is in $\overline{L}$. Therefore (c) if $\bar{x}$ and $\bar{y}$ are fully independent in $\overline{L}$, they must belong to the same component of $\overline{L}$.

[4] S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932, see p. 182.

[5] The use of the set $\overline{K}$ and of this argument was suggested by the referee to replace a more lengthy Heine-Borel argument.

If $\bar{x}$ and $\bar{y}$ are two linearly independent sets of n elements in X which are not fully independent, consider a set $\bar{v}$ of n elements which is fully independent of $\bar{x}$ and also of $\bar{y}$. Such a set $\bar{v}$ exists because $\mathfrak{X}$ is assumed to be infinite-dimensional. The subspace F can be taken to contain all the vectors x_i, y_i, and v_i. For the $\overline{L}$ which corresponds to this F, $\bar{x}$, $\bar{y}$, and $\bar{v}$ must lie in the same component by (c). Then by (b) there exists a $U \in \mathfrak{G}$ with $\overline{U}(\bar{x}) = \bar{y}$. This completes the proof.

Next we consider all our transformations as being in the space $\mathfrak{E}(\mathfrak{X})$. For $\mathfrak{E}(\mathfrak{X})$ the strong topology is defined, following the ideas of von Neumann,[6] as that where the neighborhoods of a transformation U_0 are those of the form

$$N(U_0; x_1, \cdots, x_k, \epsilon) = \{T \in \mathfrak{E}(\mathfrak{X}) \mid \|T(x_i) - U_0(x_i)\| < \epsilon, i = 1, \cdots, k\}$$

where $\epsilon > 0$ and $x_i \in \mathfrak{X}$, $i = 1, \cdots, k$.

THEOREM 2. *A subset $\mathfrak{A}$ of $\mathfrak{E}(\mathfrak{X})$ satisfying the conclusion of Theorem 1 is dense in the strong topology of $\mathfrak{E}(\mathfrak{X})$.*

It suffices to show that the neighborhood $N(U_0; x_1, \cdots, x_k, \epsilon)$ contains an operator in $\mathfrak{A}$. By renumbering, if necessary, we let $x_1, \cdots, x_r$ be a linearly independent subset of the x's which generates the same linear manifold as is generated by all of them. We put $x_i = \sum_{j=1}^{r} a_{ij} x_j$, $i = r+1, \cdots, k$ and $A = \max |a_{ij}|$ for these values of i and for $j = 1, \cdots, r$. Also we set $\eta = \min \{\epsilon, \epsilon/(rA)\}$. We choose $w_1, \cdots, w_r$ in $\mathfrak{X}$ where each w_i is linearly independent of the elements $U_0(x_i)$, $i = 1, \cdots, r$, and of the previously selected w's and each $\|w_i\| < \eta$. The collection $\{U_0(x_i) + w_i\}$ is a linearly independent set. By the assumption on $\mathfrak{A}$, there exists $T \in \mathfrak{A}$ such that $T(x_i) = U_0(x_i) + w_i$, $i = 1, \cdots, r$. Thus for these values of i, $\|T(x_i) - U_0(x_i)\| < \epsilon$. For $i = r+1, \cdots, k$ we have

$$\|T(x_i) - U_0(x_i)\| = \left\| (T - U_0)\left(\sum_{j=1}^{r} a_{ij} x_j \right) \right\| \leqq \sum_{j=1}^{r} |a_{ij}| \, \|T(x_j) - U_0(x_j)\| < \epsilon.$$

Then $T \in N(U_0; x_1, \cdots, x_k; \epsilon)$.

CORNELL UNIVERSITY

[6] J. von Neumann, *Zur Algebra der Funktionaloperatoren und Theorie des normalen Operatoren*, Math. Ann. vol. 102 (1929) pp. 370–427.

ON LACUNARY DIRICHLET SERIES

I. I. HIRSCHMAN, JR., AND J. A. JENKINS

The following theorem is suggested by a result of S. Mandelbrojt [2, p. 101][1] concerning lacunary Fourier series.

THEOREM 1. *Let* $f(s)=\sum_{k=1}^{\infty} a_k e^{-\lambda_k s}$, *where* $0<\lambda_1<\lambda_2<\lambda_3<\cdots$, $\lim_{k\to\infty}\lambda_k=\infty$. *We denote by* γ_c *the abscissa of convergence of this series and by* γ_a *the abscissa of analyticity. It is assumed that* $\gamma_c<\infty$, $\gamma_a>-\infty$. *Let* $\nu<1$ *be the exponent of convergence of the* λ_k's *and suppose that*

$$|f(s_a+\sigma)| = O[e^{-(1/\sigma)^{\mu}}] \qquad (\sigma\to 0+)$$

where $s_a=\gamma_a+i\tau_a$. *Then if* $\mu>\nu/(1-\nu)$, $f(s)\equiv 0$.

We shall in what follows prove a more general theorem including Theorem 1 as a special case. The methods of the present paper are closely related to and in part derived from the work of L. Schwartz [4]. However no appeal is made to other than standard theorems of analysis.

Let $m(0)=0$ and let $m(\sigma)$ be an increasing function defined for $0\leqq\sigma\leqq a$, $a>0$. A function $f(s)$ which we may suppose analytic in the half-plane $\sigma>\gamma_1$ is said to have a zero of modular order $m(\sigma)$ at $s_1=\gamma_1+i\tau_1$ if

$$|f(s_1+\sigma)| \leqq m(\sigma) \qquad (0<\sigma\leqq b)$$

for some $b>0$. Let us define σ as a function of ρ, $\sigma=\eta(\rho)$, by the equation $e^{-\rho\sigma}=m(\sigma)$. Since $m(\sigma)$ decreases as σ decreases to 0 this definition is effective. It is $\eta(\rho)$ which we shall use as the measure of the zero of $f(s)$. As an example, if $m(\sigma)=\exp[-(\sigma^{-\mu})]$ then $\eta(\rho)=\rho^{-1/(\mu+1)}$.

As the measure of the degree of lacunarity of our Dirichlet series we introduce

$$\zeta(\rho)=\sum_{k=1}^{\infty}\log\left(1+\frac{2\rho}{\lambda_k}\right).$$

If the sequence $\{\lambda_k\}_1^{\infty}$ has exponent of convergence ν, then by a standard theorem on integral functions [5, p. 251],

$$\zeta(\rho)=O(\rho^{\nu+\epsilon}) \qquad (\rho\to+\infty).$$

Presented to the Society, September 2, 1949; received by the editors April 11, 1949 and, in revised form, May 20, 1949.

[1] Numbers in brackets refer to the references cited at the end of the paper.

We may now state our principal theorem.

THEOREM 2. *Let* $0<\lambda_1<\lambda_2<\cdots$, $\sum_1^\infty \lambda_k^{-1}<\infty$, *and let* $\zeta(\rho)$ *be defined as above. Let*

$$f(s) = \sum_{k=1}^{\infty} a_k e^{-\lambda_k s}$$

have abscissa of convergence $\gamma_c<\infty$ *and abscissa of analyticity* $\gamma_a>-\infty$, *and let* $\gamma_a+i\tau_a$ *be a zero of* $f(s)$ *of modular order* $m(\sigma)$. *Then if* $\eta(\rho)$ *is defined as above and if*

$$\liminf_{\rho\to\infty} \zeta(\rho) - \rho\eta(\rho) = -\infty, \tag{1}$$

we have $f(s)\equiv 0$.

In the case considered in Theorem 1 we have, as we have seen, $\zeta(\rho)=O(\rho^{\nu+\epsilon})$, $\eta(\rho)=\rho^{-1/(\mu+1)}$. The assumption $\mu>\nu/(1-\nu)$ shows that condition (1) is satisfied. Thus Theorem 2 does include Theorem 1.

It is clearly sufficient to prove our theorem for $\gamma_a=\tau_a=0$.

It is immediately verifiable that the function

$$\frac{1}{(1+z)} \prod_{k=1,\,k\neq n}^{\infty} \left(\frac{\lambda_k+\rho-z}{\lambda_k+\rho+z}\right) = k_n(z,\rho) \qquad (\rho>0)$$

is analytic for $x\geqq 0$, and that

$$\int_{-\infty}^{\infty} |\, k_n(x+iy,\rho)\,|^2 dy \leqq \pi \qquad (x\geqq 0).$$

By a simple extension of Plancherel's theorem, see [3, p. 8], we see that if

$$\phi(n,\rho,\sigma) = \operatorname*{l.i.m.}_{T\to\infty}{}^{(2)} \frac{1}{2\pi i}\int_{-iT}^{iT} \frac{k_n(z,\rho)}{k_n(\lambda_n+\rho,\rho)} e^{\sigma z} dz \qquad (0\leqq\sigma<\infty),$$

then

$$\int_0^\infty \phi(n,\rho,\sigma) e^{-\sigma s} d\sigma = \frac{k_n(z,\rho)}{k_n(\lambda_n+\rho,\rho)} \qquad (x>0). \tag{2}$$

Further

$$\|\phi(n,\rho,\sigma)\|_2 \equiv \left[\int_0^\infty |\,\phi(n,\rho,\sigma)\,|^2 d\sigma\right]^{1/2} \leqq [2^{1/2} k_n(\lambda_n+\rho,\rho)]^{-1}.$$

We assert that

$$\limsup_{\rho\to\infty} \|\phi(n, \rho, \sigma)\|_2 e^{-\gamma(\rho)} < \infty. \tag{3}$$

We have

$$\begin{aligned}
&\|\phi(n, \rho, \sigma)\|_2 e^{-\gamma(\rho)} \\
&= \left[\frac{1}{2^{1/2}} \prod_{k=1, k\neq n}^{\infty} \left(\frac{\lambda_n + \lambda_k + 2\rho}{\lambda_k - \lambda_n}\right)\right] [1 + \lambda_n + \rho] \left[\prod_{k=1}^{\infty} \left(\frac{\lambda_k}{\lambda_k + 2\rho}\right)\right] \\
&= \frac{1}{2^{1/2}} \left[\prod_{k=1, k\neq n}^{\infty} \left(1 + \frac{\lambda_n}{\lambda_k + 2\rho}\right)\right] \left[\frac{1 + \lambda_n + \rho}{\lambda_n + 2\rho} \lambda_n\right] \\
&\quad \cdot \left[\prod_{k=1, k\neq n}^{\infty} \left(1 - \frac{\lambda_n}{\lambda_k}\right)\right]^{-1} \\
&\sim \frac{\lambda_n}{2^{3/2}} \left[\prod_{k=1, k\neq n}^{\infty} \left(1 - \frac{\lambda_n}{\lambda_k}\right)\right]^{-1} \qquad (\rho \to \infty),
\end{aligned}$$

which proves our assertion.

We define

$$F(\rho, s) = e^{-\rho s} f(s).$$

We assert that if

$$\|F(\rho, \sigma)\|_2 \equiv \left[\int_0^{\infty} |F(\rho, \sigma)|^2 d\sigma\right]^{1/2},$$

then

$$\limsup_{\rho\to\infty} \|F(\rho, \sigma)\|_2 e^{\rho\eta(\rho)} < \infty. \tag{4}$$

Let M be a bound for $f(\sigma)$ for $0 < \sigma < \infty$. We have

$$\begin{aligned}
\|F(\rho, \sigma)\|_2 &\leqq \left[\int_0^{\eta(\rho)} e^{-2\rho\sigma} |f(\sigma)|^2 d\sigma\right]^{1/2} + \left[\int_{\eta(\rho)}^{\infty} e^{-2\rho\sigma} |f(\sigma)|^2 d\sigma\right]^{1/2} \\
&\leqq m(\eta(\rho)) + M e^{-\rho\eta(\rho)} \\
&\leqq (M + 1) e^{-\rho\eta(\rho)},
\end{aligned}$$

which proves relation (4).

It follows from a well known theorem concerning restricted overconvergence of Dirichlet series, see [1, p. 141], that there exists an increasing sequence of integers l_k for which

$$\lim_{k\to\infty} \sum_{j=0}^{l_k} a_j e^{-\lambda_j \sigma} = f(\sigma)$$

uniformly for $\epsilon \leqq \sigma < \infty$ for any $\epsilon > 0$. Thus if $\epsilon > 0$,

$$\int_0^\infty \phi(n, \rho, \sigma) F(\sigma + \epsilon, \rho) d\sigma$$

$$= \lim_{k \to \infty} \int_0^\infty \phi(n, \rho, \sigma) e^{-\rho(\sigma+\epsilon)} \sum_{j=1}^{l_k} a_j e^{-\lambda_j(\sigma+\epsilon)} d\sigma.$$

By equation (2)

$$\int_0^\infty \phi(n, \rho, \sigma) e^{-\rho(\sigma+\epsilon)} a_j e^{-\lambda_j(\sigma+\epsilon)} d\sigma = \begin{cases} 0, & j \neq n, \\ e^{-\epsilon\rho - \epsilon\lambda_n} a_n, & j = n. \end{cases}$$

Hence

$$\int_0^\infty \phi(n, \rho, \sigma) F(\rho, \sigma + \epsilon) d\sigma = e^{-\epsilon\rho - \epsilon\lambda_n} a_n.$$

By Schwarz's inequality

$$|a_n| \leqq e^{\epsilon\rho + \epsilon\lambda_n} \|\phi(n, \rho, \sigma)\|_2 \|F(\sigma + \epsilon, \rho)\|_2.$$

If ϵ is allowed to approach zero through positive values, we obtain

$$|a_n| \leqq \|\phi(n, \rho, \sigma)\|_2 \|F(\sigma, \rho)\|_2.$$

Using equations (3) and (4) we see that

$$\log |a_n| \leqq \zeta(\rho) - \rho\eta(\rho) + C$$

where C is a constant which depends on n but not upon ρ. If we allow ρ to increase without limit, assumption (1) of Theorem 2 implies that

$$\log |a_n| = -\infty, \qquad \text{that is, } a_n = 0.$$

Since this holds for each $n = 1, 2, \cdots, f(s) \equiv 0$ and our theorem is proved.

We shall now prove that if assumption (1) is replaced by the weaker assumption

$$\liminf_{\rho \to \infty} c\zeta(\rho) - \rho\eta(\rho) = -\infty$$

where $c < 2^{-3/2}$, then Theorem 2 is false.

Let us take for the constants λ_k, $0 < a < b < 1 < 4 < 9 < \cdots$. As in Titchmarsh [5, p. 271] we find that

$$\log \left[\prod_1^\infty \left(1 + \frac{x}{k^2}\right) \right] \sim \pi x^{1/2} \qquad (x \to \infty). \tag{5}$$

It follows that

$$\zeta(\rho) = \log\left(1+\frac{2\rho}{a}\right) + \log\left(1+\frac{2\rho}{b}\right) + \sum_{k=1}^{\infty} \log\left(1+\frac{2\rho}{k^2}\right) \sim \pi(2\rho)^{1/2} \qquad (\rho \to \infty). \tag{6}$$

We define

$$f(s) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{e^{sz}}{\left(1+\frac{z}{a}\right)\left(1+\frac{z}{b}\right)\prod_{k=1}^{\infty}\left(1+\frac{z}{k^2}\right)}\,dz. \tag{7}$$

Deforming the line of integration of this integral to $\mathrm{Rl}\, z = -(n+1/2)^2$ and allowing n to approach infinity we obtain for s real and positive an expansion of $f(s)$ into a Dirichlet series with exponents a, b, 1, 4, 9, $\cdots$ which then converges in the half-plane $\mathrm{Rl}\, s>0$. The details of this are left to the reader. Again deforming the line of integration, this time to $\mathrm{Rl}\, z=(\pi/2\sigma)^2$, and using equation (5), we see that if $\epsilon>0$ is fixed, then for $\sigma>0$ sufficiently small we have

$$|f(\sigma)| \leqq \exp\left[-\frac{1-2\epsilon}{4}\frac{\pi^2}{\sigma}\right] M(\sigma)$$

where

$$M(\sigma) = \frac{1}{2\pi}\int_{\mathrm{Rl}\, z=(\pi/2\sigma)^2} \left|\frac{dz}{\left(1+\frac{z}{a}\right)\left(1+\frac{z}{b}\right)}\right| = o(1) \qquad (\sigma \to 0+).$$

Thus we may associate with $f(s)$ and the origin the modular order $m(\sigma)=\exp\left[-((1-2\epsilon)/4)(\pi^2/\sigma)\right]$. We immediately find that

$$\eta(\rho) = \frac{\pi}{2}(1-2\epsilon)^{1/2}\rho^{-1/2}. \tag{8}$$

Comparing equations (6) and (8), we deduce that Theorem 2 is false if $c<2^{-3/2}$, which is what we wished to show.

References

1. V. Bernstein, *Leçons sur les progrès récents de la théorie des series de Dirichlet*, Paris, 1933.

2. S. Mandelbrojt, *Analytic functions and classes of infinitely differentiable func-*

tions, The Rice Institute Pamphlet, vol. 29, 1942.
3. R. Paley and N. Wiener, *Fourier transforms in the complex domain*, New York, 1934.
4. L. Schwartz, *Études des sommes d'exponentielles réelles*, Paris, 1943.
5. E. Titchmarsh, *The theory of functions*, 2d ed., Oxford, 1939.

HARVARD UNIVERSITY

ON THE ABSOLUTE CONVERGENCE OF TRIGONOMETRICAL SERIES

TAMOTSU TSUCHIKURA AND SHIGEKI YANO

1. THEOREM 1. *Suppose that a trigonometrical series*

$$\sum \rho_n \cos (nx - \alpha_n) \qquad (\rho_n \geqq 0,\ n = 1, 2, \cdots)$$

and its conjugate series

$$\sum \rho_n \sin (nx - \alpha_n)$$

are convergent absolutely at $x = x_0$ and $x = x_1$, respectively. If

$$x_1 - x_0 = p\pi/q \qquad (p/q \text{ irreducible})$$

where p is an integer positive, negative, or zero, and q is an odd integer, then

$$\sum \rho_n < \infty.$$

We shall see in Theorem 2 that the above theorem is no longer true if p/q is replaced by p'/q' with an even q' and $p' \neq 0$, or by an irrational number.

PROOF OF THEOREM 1. If we put $x_1 - x_0 = p\pi/q \equiv h$, then by Fatou's theorem[1]

$$\sum \rho_n \left| \cos (n(x_1 - sh) - \alpha_n) \right| < \infty \tag{1}$$

for every odd s. Hence from the identity

$$\sin (nx_1 - \alpha_n) = \cos nsh \sin (n(x_1 - sh) - \alpha_n) + \sin nsh \cos (n(x_1 - sh) - \alpha_n)$$

we deduce immediately that

$$\sum \rho_n \left| \cos nsh \sin (n(x_1 - sh)) - \alpha_n) \right| < \infty. \tag{2}$$

Received by the editors October 18, 1948 and, in revised form, May 14, 1949.

[1] See, for example, A. Zygmund, *Trigonometrical series*, p. 134.

On the other hand from (1) we have

$$\sum \rho_n \left| \cos nsh \cos \left(n(x_1 - sh) - \alpha_n\right) \right| < \infty, \tag{3}$$

hence from (2) and (3)

$$\sum \rho_n \left| \cos nsh \right| < \infty \tag{4}$$

if s is an odd number.

If we replace s in (4) by an odd q, then we get

$$\sum \rho_n = \sum \rho_n \left| \cos nq \cdot \frac{p}{q} \pi \right| < \infty$$

which is the required inequality.

We may also prove the theorem geometrically by using the cited Fatou's theorem concerning the symmetry of points of absolute convergence, and the fact that if the series

$$\sum \rho_n \cos (nx - \alpha_n)$$

converges absolutely at a point x', then also it converges at the point $x'+\pi$.

THEOREM 2. *For a given sequence* $\{\rho_\nu\}$ *of positive numbers such that* $\sum \rho_\nu = \infty$, *and an irreducible fraction* p'/q' $(\neq 0)$ *with an even* q' *or an irrational number* λ, *we can construct a trigonometrical series*

$$\sum \rho_\nu \cos n_\nu x$$

which is absolutely convergent at $x = p' \cdot \pi/q'$ *or* $x = \lambda\pi$ *respectively.*

Evidently the conjugate series $\sum \rho_\nu \sin n_\nu x$ converges absolutely at $x = 0$.

For the proof of the first part of the theorem, we put

$$n_\nu \equiv (2\nu + 1)q'/2 \qquad (\nu = 1, 2, \cdots);$$

then, p' being an odd number, we have

$$\sum \rho_\nu \left| \cos n_\nu \frac{p'}{q'} \pi \right| = \sum \rho_\nu \left| \cos (2\nu + 1)p' \frac{\pi}{2} \right| = 0.$$

For the second part we consider the sequence

$$\{n\lambda\pi\}_{n=1,2,\cdots}$$

which is uniformly distributed in the interval $(0, 2\pi)$ by Weyl's theorem; then we can find a strictly increasing sequence $\{n_\nu\}$ such that

$$|\cos n_\nu \lambda\pi| \leqq 1 \Big/ \Big(\sum_{\mu=1}^{\nu} \rho_\mu\Big)^2.$$

Thus we have

$$\sum_\nu \rho_\nu |\cos n_\nu \lambda\pi| \leqq \sum_\nu \frac{\rho_\nu}{\Big(\sum_{\mu=1}^{\nu} \rho_\mu\Big)^2} < \infty.^2$$

THEOREM 3. *Let* $\sum \rho_n = \infty$, *and one of the following conditions be satisfied*:

(i) ρ_n *is non-increasing,*

(ii) ρ_n/ρ_m *is bounded uniformly for* $n \geqq m > 0$,

(iii) $\sum_{\nu=1}^{n} 1/\rho_\nu = O(n^2)$.

If $\sum \rho_n \cos(nx - \alpha_n)$ *converges absolutely at* $x = x_0$, *then it does not converge elsewhere; moreover the conjugate series* $\sum \rho_n \sin(nx - \alpha_n)$ *does not converge anywhere.*

The first part of the theorem is due to R. Salem.[3] For the proof of the second part we suppose, if possible, that

$$\sum \rho_n \sin(nx - \alpha_n)$$

converges absolutely at $x = x_1$, then by (4)

$$\sum \rho_n |\cos n(x_0 - x_1)| < \infty$$

from which we deduce that

$$\sum \rho_n < \infty$$

by the argument used by Salem to prove the first part. This is a contradiction.

2. Consider the trigonometrical series

$$\sum \rho_n \cos(nx - \alpha_n) \qquad (\rho_n \geqq 0;\ n = 1, 2, \cdots). \tag{5}$$

If $\sum \rho_n = \infty$, then the set of absolute convergence points of (5) is of measure zero by the Denjoy-Lusin theorem. On the other hand by Theorem 3, (iii), if $\sum_{\nu=1}^{n} 1/\rho_\nu = O(n^2)$, the series (5) converges ab-

[2] This proof is shown by Professor Izumi. We can easily find a sequence $\{n_\nu\}$ and a set E of generalized Cantor type such that $\lambda\pi \in E$ and $|\cos n_\nu| < 1/(\sum_{\mu=1}^{\nu} \rho_\mu)^2$ $(\nu = 1, 2, \cdots)$ for every $x \in E$, that is $\sum \rho_\nu |\cos n_\nu x| < \infty$, for every x in E which includes $\lambda\pi$. Therefore we see that each of the series $\sum \rho_\nu \cos n_\nu x$ and $\sum \rho_\nu \sin n_\nu x$ has the perfect set of absolute convergence points.

[3] R. Salem, *The absolute convergence of trigonometrical series*, Duke Math. J. vol. 8 (1941).

solutely at most at one point. As a counterpart of this theorem we prove the following theorem.

THEOREM 4. *If* $\sum \rho_n = \infty$, *and* $\rho_n = O(1/n)$, *then the set* E *of absolute convergence points of* (5) *is of* α*-capacity* $(0<\alpha<1)$ *zero.*[4]

For the proof, let E be of positive α-capacity $(0<\alpha<1)$; then by the Salem-Zygmund theorem,[5] there exists positive distribution $\mu \in \mathrm{Lip}\ \alpha$ concentrated on E. If we put

$$d\mu \sim \frac{1}{2\pi} + \sum (a_n \cos nx + b_n \sin nx),$$

then

$$\mu(x) - \frac{x}{2\pi} + C \sim \sum_1^\infty \frac{a_n \sin nx - b_n \cos nx}{n}$$

and, $\mu(x)$ being monotone, the left-hand side is of bounded variation. Hence by a theorem due to Zygmund[6] we have

$$\sum (|a_n| + |b_n|)/n < \infty. \tag{6}$$

Let us put $\mu(H) \equiv \int_H d\mu$ for $H \subseteq (0, 2\pi)$. By Egoroff's theorem, for any $\epsilon > 0$ $(\epsilon < 1/6)$, there exist a set $\mathcal{E} \subseteq E$ and a positive number M such that $\mu(\mathcal{E}) > 1-\epsilon$ and

$$\sum_1^\infty \rho_n |\cos (nx - \alpha_n)| < M \qquad (x \in \mathcal{E}).$$

Integrating both sides we get

$$\sum_{n=1}^\infty \rho_n \int_{\mathcal{E}} |\cos (nx - \alpha_n)|\, d\mu < M\mu(\mathcal{E}),$$

where

---

[4] R. Salem and A. Zygmund, *Capacity of sets and Fourier series*, Trans. Amer. Math. Soc. vol. 59 (1946) pp. 23–41. The α-capacity of sets is defined as follows: The distribution of positive mass on the unit circle $|z| = 1$, $z = re^{i\alpha}$, shall be supposed to be a non-negative additive function of sets, from which we define a point function $\mu(x)$ non-decreasing for $0 \leq x \leq 2\pi$. It will be supposed that the total distribution $\mu(2\pi) - \mu(0) = 1$. If for a (Borel) set E, $\int_E d\mu = \int_0^{2\pi} d\mu$, we say that a distribution μ is concentrated on E. A set E is said to be of positive α-capacity, if there exists a distribution μ concentrated on E and $V \equiv \int_0^{2\pi} d\mu(t)/|e^{i\alpha} - re^{it}|^\alpha$ is uniformly bounded in x as $r \to 1$.

[5] Salem-Zygmund, loc. cit. footnote 4.

[6] Zygmund, loc. cit. footnote 1, p. 136.

$$
\begin{aligned}
&2\int_{\mathcal{E}} |\cos(nx-\alpha_n)|\,d\mu \\
&\geqq \int_{\mathcal{E}} (1-2\cos 2\alpha_n \cos 2nx+\sin 2\alpha_n \sin 2nx)\,d\mu \\
&\geqq 1-\epsilon+\cos 2\alpha_n \int_{\mathcal{E}} \cos 2nx\,d\mu+\sin 2\alpha_n \int_{\mathcal{E}} \sin 2nx\,d\mu \\
&= 1-\epsilon+\cos 2\alpha_n \int_{E} \cos 2nx\,d\mu+\sin 2\alpha_n \int_{E} \sin 2nx\,d\mu \\
&\quad -\cos 2\alpha_n \int_{E-\mathcal{E}} \cos 2nx\,d\mu-\sin 2\alpha_n \int_{E-\mathcal{E}} \sin 2nx\,d\mu \\
&\geqq 1-\epsilon+\cos 2\alpha_n \int_0^{2\pi} \cos 2nx\,d\mu+\sin 2\alpha_n \int_0^{2\pi} \sin 2nx\,d\mu-2\epsilon \\
&\geqq 1-3\epsilon-|a_{2n}|-|b_{2n}|.
\end{aligned}
$$

Consequently

$$
2M \geqq \sum \rho_n(1-3\epsilon)-\sum \rho_n(|a_{2n}|+|b_{2n}|). \tag{7}
$$

On the other hand, by (2) and $\rho_n=O(1/n)$, $\sum\rho_n(|a_{2n}|+|b_{2n}|)$ $=O(\sum(|a_{2n}|+|b_{2n}|)/n)<\infty$ which contradicts (7).

THEOREM 5. *If* (5) *is convergent absolutely on a set of positive* α*-capacity* $(0<\alpha<1)$, *we can divide the suffixes into two classes* $\{n_i\}$ and $\{m_i\}$ *such that* $\sum 1/m_i<\infty$, $\sum\rho_{n_i}<\infty$.[7]

For, we get (6) by the hypothesis, and then we can deduce the theorem by the usual argument.

REMARK. Let μ be a distribution concentrated on E. If

$$
\int_0^{2\pi} \log\frac{1}{|e^{it}-re^{ix}|}\,d\mu(t)
$$

is uniformly bounded in x as $r\to 1$, we say that E is of positive logarithmic capacity. Salem[8] gave an example such that $\sum\rho_n=\infty$ and $\sum\rho_n\cos(nx-\alpha_n)$ converges absolutely on a set of positive logarithmic capacity. But for α-capacity such an example seems to be open.

TÔHOKU UNIVERSITY

[7] For example, see Zygmund, *Sur la caractère de convergence des séries orthogonales*, Matematica vol. 9 (1935) pp. 86–88.

[8] Salem, loc. cit. footnote 3.

ON DOUBLE FOURIER SERIES

GEN-ICHIRÔ SUNOUCHI

1. In the present note, the author proves some inequality theorems on double Fourier series. We begin with some preliminary lemmas. Let $f(x, y)$ be an integrable function and let its Fourier series be

$$f(x, y) \sim \sum_{m,n} A_{m,n}(x, y)$$

and let

$$\Delta_{\mu,\nu}(x, y) = \sum_{m=2^\mu}^{2^{\mu+1}-1} \sum_{n=2^\nu}^{2^{\nu+1}-1} A_{m,n}(x, y).$$

LEMMA 1. *Let $s_{\nu;m,n}(x, y)$ be the m, nth partial sum of Fourier series of $f_\nu(x, y)$, then*

$$\int_0^{2\pi}\int_0^{2\pi}\left(\sum_\nu {}^2 s_{\nu;m,n}(x, y)\right)^{\gamma/2} dxdy \leqq A_\gamma \int_0^{2\pi}\int_0^{2\pi}\left(\sum_\nu f_\nu^2\right)^{\gamma/2} dxdy, \qquad \gamma > 1.$$

LEMMA 2. *If $\gamma > 1$,*

$$\int_0^{2\pi}\int_0^{2\pi}\left(\sum_{\mu,\nu} |\Delta_{\mu,\nu}(x, y)|^2\right)^{\gamma/2} dxdy \leqq B_\gamma \int_0^{2\pi}\int_0^{2\pi} |f(x, y)|^\gamma dxdy \leqq B'_\gamma \int_0^{2\pi}\int_0^{2\pi}\left(\sum_{\mu,\nu} |\Delta_{\mu,\nu}(x, y)|^2\right)^{\gamma/2} dxdy.$$

These lemmas are due to J. Marcinkiewicz [2].[1]

2. THEOREM 1. *If $\gamma > 1$,*

$$\int_0^{2\pi}\int_0^{2\pi}\left(\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} |s_{2^m,2^n}(x, y) - \sigma_{2^m,2^n}(x, y)|^2\right)^{\gamma/2} dxdy \leqq C_\gamma \int_0^{2\pi}\int_0^{2\pi} |f(x, y)|^\gamma dxdy,$$

Received by the editors October 18, 1948 and, in revised form, May 14, 1949.

[1] Numbers in brackets refer to the references cited at the end of the paper.

and

$$\int_0^{2\pi}\int_0^{2\pi}\left|\max_{m,n} s_{2^m,2^n}(x,y)\right|^\gamma dxdy \leqq D_\gamma \int_0^{2\pi}\int_0^{2\pi} |f(x,y)|^\gamma dxdy.$$

Accordingly $s_{2^m,2^n}(x, y)$ *converges to* $f(x, y)$ *almost everywhere, when* $f(x, y)\in L^\gamma$ $(\gamma>1)$.

By Schwarz's inequality

$$\sum_m^\infty\sum_n^\infty |s_{2^m,2^n}-\sigma_{2^m,2^n}|^2$$

$$\leqq \sum_m^\infty\sum_n^\infty \frac{1}{2^{2m}\cdot 2^{2n}}\left(\sum_{\mu=0}^{2}\sum_{\nu=0}^{m_2}{}^{n} |s_{2^m,2^n}-s_{\mu,\nu}|\right)^2$$

$$\leqq \sum_m\sum_n \frac{1}{2^m}\frac{1}{2^n}\left(\sum_{\mu=0}^{2^m}\sum_{\nu=0}^{2^n} |s_{2^m,2^n}-s_{\mu,\nu}|^2\right)$$

$$\leqq \sum_m\sum_n \frac{1}{2^m}\frac{1}{2^n}\left\{\sum_{i=0}^{m}\sum_{j=0}^{n}\left(\sum_{\mu=2^{i-1}+1}^{2^i}\ \sum_{\nu=2^{j-1}+1}^{2^j} |s_{2^m,2^n}-s_{\mu,\nu}|^2\right)\right\}.$$

In view of Lemma 1, we get

$$\int_0^{2\pi}\int_0^{2\pi}\left(\sum_m\sum_n |s_{2^m,2^n}-\sigma_{2^m,2^n}|^2\right)^{\gamma/2} dxdy$$

$$\leqq E_\gamma\int_0^{2\pi}\int_0^{2\pi}\left[\sum_m\sum_n \frac{1}{2^m}\frac{1}{2^n}\cdot\left\{\sum_{i=0}^{m}\sum_{j=0}^{n}(2^i2^j|s_{2^m,2^n}-s_{2^{i-1},2^{j-1}}|^2)\right\}\right]^{\gamma/2} dxdy.$$

Since $|s_{2^m,2^n}-s_{2^{i-1},2^{j-1}}|\leqq\sum_{k=i}^{m}\sum_{l=j}^{n}|\Delta_{k,l}|$, we have

$$\sum_m\sum_n \frac{1}{2^m}\frac{1}{2^n}\sum_{i=0}^{m}\sum_{j=0}^{n}2^i2^j|s_{2^m,2^n}-s_{2^{i-1},2^{j-1}}|^2$$

$$\leqq \sum_m\sum_n \frac{1}{2^m}\frac{1}{2^n}\sum_{i=0}^{m}\sum_{j=0}^{n}2^i2^j\left[\sum_{k=i}^{m}\sum_{l=j}^{n}\left\{|\Delta_{k,l}|(2^k2^l)^{1/4}(2^k2^l)^{-1/4}\right\}^2\right]$$

$$\leqq \sum_m\sum_n \frac{1}{2^m}\frac{1}{2^n}\sum_{i=0}^{m}\sum_{j=0}^{n}2^i2^j\left\{\sum_{k=i}^{m}\sum_{l=j}^{n}|\Delta_{k,l}|^2(2^k2^l)^{1/2}\right\}$$

$$\cdot\left\{\sum_{k=i}^{\infty}\sum_{l=j}^{\infty}(2^k2^l)^{-1/2}\right\}$$

$$\leqq \sum_m \sum_n \frac{1}{2^m}\frac{1}{2^n}\sum_{i=0}^{m}\sum_{j=0}^{n}(2^i 2^j)^{1/2}\sum_{k=i}^{m}\sum_{l=j}^{n}|\Delta_{k,l}|^2(2^k 2^l)^{1/2}$$

$$\leqq \sum_m \sum_n \frac{1}{2^m}\frac{1}{2^n}\sum_{k=0}^{m}\sum_{l=0}^{n}|\Delta_{k,l}|^2(2^k 2^l)^{1/2}\sum_{i=0}^{k}\sum_{j=0}^{l}(2^i 2^j)^{1/2}$$

$$\leqq \sum_m \sum_n \frac{1}{2^m}\frac{1}{2^n}\sum_{k=0}^{m}\sum_{l=0}^{n}|\Delta_{k,l}|^2 2^k 2^l$$

$$\leqq \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}|\Delta_{k,l}|^2 2^k 2^l \sum_{m=k}^{\infty}\sum_{n=l}^{\infty}\frac{1}{2^m}\frac{1}{2^n}$$

$$\leqq \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}|\Delta_{k,l}|^2.$$

By using Lemma 2, we get the first half of the theorem. The remaining part is an easy consequence of the former.

3. LEMMA 3. *If $k>1$ and $\gamma>1$, we get*

$$\int_0^{2\pi}\int_0^{2\pi}\Big(\sum_\nu |s_{\nu;m,n}|^k\Big)^{\gamma/k}dxdy \leqq E_\gamma \int_0^{2\pi}\int_0^{2\pi}\Big(\sum_\nu |f_\nu|^k\Big)^{\gamma/k}dxdy.$$

The proof is carried out analogously to Marcinkiewicz's proof of Lemma 1, using the result of Boas and Bochner [1].

THEOREM 2. *If $k\geqq 2$ and $r>1$,*

$$\int_0^{2\pi}\int_0^{2\pi}\Big(\sum_m\sum_n \frac{|s_{m,n}-\sigma_{m,n}|^k}{mn}\Big)^{\gamma/k}dxdy \leqq F_\gamma \int_0^{2\pi}\int_0^{2\pi}|f(x,y)|^\gamma dxdy.$$

Accordingly if $f(x, y)\in L^\gamma$ $(\gamma>1)$,

$$\Big(\sum_{\mu=1}^{m}\sum_{\nu=1}^{n}|s_{\mu,\nu}-f|^k\Big) = o(mn) \qquad (k\geqq 1)$$

almost everywhere.

The proof is analogous to my previous paper [4], using Lemma 3. The last part of the theorem was obtained by Marcinkiewicz [3] in more generalized form.

LEMMA 4. *If* $r>1$, *then*

$$\int_0^{2\pi}\int_0^{2\pi}\left(\sum_\mu\sum_\nu|\Delta_{\mu,\nu}|^2\right)^{\gamma/2}dxdy \leqq \int_0^{2\pi}\int_0^{2\pi}\left(\sum_m\sum_n\frac{|s_{m,n}-\sigma_{m,n}|^2}{mn}\right)^{\gamma/2}dxdy.$$

The proof is easy (cf. the author's note [5]).

THEOREM 3. *If* $\gamma>1$,

$$\int_0^{2\pi}\int_0^{2\pi}|f(x,y)|^\gamma dxdy \leqq G_\gamma\int_0^{2\pi}\int_0^{2\pi}\left(\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}|s_{2^m,2^n}-\sigma_{2^m,2^n}|^2\right)^{\gamma/2}dxdy$$

and

$$\int_0^{2\pi}\int_0^{2\pi}|f(x,y)|^\gamma dxdy \leqq H_\gamma\int_0^{2\pi}\int_0^{2\pi}\left(\sum_m\sum_n\frac{|s_{m,n}-\sigma_{m,n}|^2}{mn}\right)^{\gamma/2}dxdy.$$

The proof is immediate by Lemma 4 and Lemma 2.

BIBLIOGRAPHY

1. R. P. Boas, Jr., and S. Bochner, *On a theorem of M. Riesz for Fourier series*, J. London Math. Soc. vol. 14 (1939) pp. 62–73.

2. J. Marcinkiewicz, *Sur les multiplicateurs des séries de Fourier*, Studia Mathematica vol. 8 (1939) pp. 78–91.

3. ———, *Sur la sommabilité H_k de séries de Fourier*, C. R. Acad. Sci. Paris vol. 208 (1939) pp. 782–784.

4. G. Sunouchi, *On the strong summability of Fourier series*, Proc. Imp. Acad. Tokyo vol. 19 (1943) pp. 420–423.

5. ———, *On Walsh-Kaczmarz series*, to be published in the Proceedings of the American Mathematical Society.

TOHOKU IMPERIAL UNIVERSITY

ON THE STRONG SUMMABILITY OF FOURIER SERIES

GEN-ICHIRÔ SUNOUCHI

This paper consists of three parts. In the first part we complete one of the deepest theorems due to Littlewood and Paley [3, 4].[1] Our theorem was already proved by Marcinkiewicz and Zygmund [4]. But the proof given here is very simple and direct. In the second part a strong summability theorem is proved. It is the completion of our former theorem [5]. We use the method due to Zygmund [8] for the proof. Finally we prove a strong summability theorem concerning lacunary sequences of partial sums (Zalcwasser [6]).

1. Let $f(\theta)$ be an integrable function with period 2π and its Fourier series be $f(\theta)\sim\sum_{n=1}^{\infty}(a_n\cos n\theta+b_n\sin n\theta)$, assuming $a_0=0$ for the sake of simplicity. If we put $z=\rho e^{i\theta}$ and

$$\phi(z)=\sum_{n=1}^{\infty}(a_n-ib_n)z^n,$$

then $f(\theta)$ becomes the boundary function of the harmonic function $\Re\phi(z)$. Let us denote by p, q, and r real numbers satisfying

$$r>1, \qquad 1<p\leqq 2\leqq q<\infty,$$

and by $A, B, \cdots$ absolute constants, not always the same from one occurrence to another.

THEOREM 1. *Let $f(\theta)$ be a function of the class L^r, then*

$$A\int_0^{2\pi}|\phi(e^{i\theta})|^r d\theta\leqq\int_0^{2\pi}\left\{\int_0^1(1-\rho)\,|\phi'(z)|^2 d\rho\right\}^{r/2}d\theta \leqq B\int_0^{2\pi}|\phi(e^{i\theta})|^r d\theta, \tag{1.1}$$

$$\int_0^{2\pi}\left\{\int_0^1(1-\rho)^{q-1}|\phi'(z)|^q d\rho\right\}^{r/q}d\theta\leqq C\int_0^{2\pi}|\phi(e^{i\theta})|^r d\theta, \tag{1.2}$$

$$D\int_0^{2\pi}|\phi(e^{i\theta})|^r d\theta\leqq\int_0^{2\pi}\left\{\int_0^1(1-\rho)^{p-1}|\phi'(z)|^p d\rho\right\}^{r/p}d\theta. \tag{1.3}$$

(1.1) is due to Littlewood and Paley [4]. In order to prove (1.2) we put

Received by the editors October 18, 1948 and, in revised form, May 14, 1949.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

$$\Phi(\theta) = \max\left(|\phi(z)|; z \in S(\theta)\right)$$

where $S(\theta)$ is a kite-shaped region.[2] Then $|\phi'| \leqq A\Phi/(1-\rho)$. If we put

$$g(\theta) = \left(\int_0^1 (1-\rho)\,|\phi'(z)|^2 d\rho\right)^{1/2},$$

then we have

$$\begin{aligned}
&\left\{\int_0^{2\pi}\left(\int_0^1 (1-\rho)^{q-1}|\phi'|^q d\rho\right)^{r/q} d\theta\right\}^{1/r} \\
&\quad \leqq \left\{\int_0^{2\pi}\left(\int_0^1 |\phi'|^2 (1-\rho) A\Phi^{q-2} d\rho\right)^{r/q} d\theta\right\}^{1/r} \\
&\quad \leqq \left\{\int_0^{2\pi} (A\Phi^{q-2}\cdot g^2)^{r/q} d\theta\right\}^{1/r} \\
&\quad \leqq \left(B\int_0^{2\pi} \Phi^{(1-2/q)r} g^{2r/q} d\theta\right)^{1/r}.
\end{aligned}$$

Since $0<2/q<1$, if we apply Hölder's inequality, the right-hand side becomes less than

$$\begin{aligned}
&C\left\{\left(\int_0^{2\pi}\Phi^r d\theta\right)^{1-2/q}\left(\int_0^{2\pi}|g|^r d\theta\right)^{2/q}\right\}^{1/r} \\
&\qquad \leqq C\left\{\left(\int_0^{2\pi}\Phi^r d\theta\right)^{1/r}\right\}^{1-2/q}\cdot\left\{\left(\int_0^{2\pi}|g|^r d\theta\right)^{1/r}\right\}^{2/q}.
\end{aligned}$$

By the maximal theorem due to Hardy and Littlewood the last term is less than

$$\begin{aligned}
&D\left\{\left(\int_0^{2\pi}|\phi|^r d\theta\right)^{1/r}\right\}^{1-2/q}\cdot\left\{\left(\int_0^{2\pi}|g|^r d\theta\right)^{1/r}\right\}^{2/q} \\
&\qquad \leqq E\left(\int_0^{2\pi}|\phi|^r d\theta\right)^{1/r}.
\end{aligned}$$

Thus we get (1.2) by (1.1).

For the proof of (1.3) we put

$$h(\theta) = \left(\int_0^1 (1-\rho)^{p-1}|\phi'(z)|^p d\rho\right)^{1/p}.$$

[2] See Littlewood and Paley [4].

Then we have

$$\left\{\int_0^{2\pi}\left(\int_0^1 (1-\rho)\,|\,\phi'(z)\,|^2 d\rho\right)^{r/2} d\theta\right\}^{1/r}$$

$$\leqq \left\{\int_0^{2\pi}\left(\int_0^1 |\,\phi'\,|^p (1-\rho)^{p-1} A\,\Phi^{2-p} d\rho\right)^{r/2} d\theta\right\}^{1/r}$$

$$\leqq \left\{\int_0^{2\pi} (A\,\Phi^{2-p}\,|\,h(\theta)\,|^p)^{r/2} d\theta\right\}^{1/r} \leqq B\left\{\int_0^{2\pi} \Phi^{(1-p/2)r}\cdot h^{pr/2} d\theta\right\}^{1/r}$$

$$\leqq C\left\{\left(\int_0^{2\pi} \Phi^r d\theta\right)^{1-p/2}\left(\int_0^{2\pi} h^r d\theta\right)^{p/2}\right\}^{1/r}$$

$$\leqq C\left\{\left(\int_0^{2\pi} \Phi^r d\theta\right)^{1/r}\right\}^{1-p/2}\left\{\left(\int_0^{2\pi} h^r d\theta\right)^{1/r}\right\}^{p/2}$$

$$\leqq D\left\{\left(\int_0^{2\pi} |\,\phi\,|^r d\theta\right)^{1/r}\right\}^{1-p/2}\left\{\left(\int_0^{2\pi} h^r d\theta\right)^{1/r}\right\}^{p/2}.$$

Thus we have

$$\left(\int_0^{2\pi} g^r d\theta\right)^{1/r} \leqq E\left(\int_0^{2\pi} h^r d\theta\right)^{1/r}.$$

Hence by (1.1) we get (1.3).

2. THEOREM 2. *Let $s_n(e^{i\theta})$ and $\tau_n(e^{i\theta})$ denote the partial sums and the arithmetic mean of Fourier power series of $\phi(e^{i\theta})$ belonging to the class H^r, respectively. Then we have*

$$A\int_0^{2\pi} |\,\phi(e^{i\theta})\,|^r d\theta \leqq \int_0^{2\pi}\left(\sum_{n=1}^{\infty} |\,s_n(e^{i\theta}) - \tau_n(e^{i\theta})\,|^2/n\right)^{r/2} d\theta$$

(2.1)

$$\leqq B\int_0^{2\pi} |\,\phi(e^{i\theta})\,|^r d\theta,$$

$$\int_0^{2\pi}\left(\sum_{n=1}^{\infty} \frac{1}{n} |\,s_n(e^{i\theta}) - \tau_n(e^{i\theta})\,|^q\right)^{r/q} d\theta \leqq C\int_0^{2\pi} |\,\phi(e^{i\theta})\,|^r d\theta, \tag{2.2}$$

$$D\int_0^{2\pi} |\,\phi(e^{i\theta})\,|^r d\theta \leqq \int_0^{2\pi}\left(\sum_{n=1}^{\infty} \frac{1}{n} |\,s_n(e^{i\theta}) - \tau_n(e^{i\theta})\,|^p\right)^{r/p} d\theta. \tag{2.3}$$

(2.1) is due to Zygmund [8, 9], and (2.2) is given in my former paper [5]. For the proof of the theorem we require the following lemmas.

LEMMA 1. *Let $f_1, f_2, \cdots$ be a sequence of integrable functions and let $s_{n,\nu}$ denote the νth partial sum of the Fourier series of f_n. Then*

$$\int_0^{2\pi} \left(\sum_{n=1}^{\infty} | s_{n,k_n}(\theta) |^m \right)^{r/m} d\theta \leqq A \int_0^{2\pi} \left(\sum_{n=1}^{\infty} | f_n |^m \right)^{r/m} d\theta,$$

where $r>1$ and $m>1$.

The inequality is also valid for conjugate series.

This lemma is due to Boas and Bochner [1].

If we write

$$s_\nu(\rho, \theta) = \sum_{n=1}^{\nu} (a_n \cos n\theta + b_n \sin n\theta)\rho^n,$$

$$f(\rho, \theta) = \sum_{n=1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta)\rho^n,$$

then we have the following lemma.

LEMMA 2. *Let $s_{n,\nu}(\rho, \theta)$ denote the sum derived from $f_n(\theta)$ similarly as $s_\nu(\rho, \theta)$ from $f(\theta)$. Then under the hypothesis of Lemma 1 we have*

$$\int_0^{2\pi} \left(\sum_{n=1}^{\infty} | s_{n,k_n}(\rho_n, \theta) |^m \right)^{2/m} d\theta \leqq A \int_0^{2\pi} \left(\sum_{n=1}^{\infty} | f_n(\theta) |^m \right)^{r/m} d\theta,$$

where $0 \leqq \rho_n \leqq 1$. The sums $s_{n,\nu}$ may be replaced by the conjugate ones.

This lemma is easily derived from Lemma 1 by Zygmund's argument by a slight modification.

LEMMA 3. *Let $0 \leqq \rho_n < 1$ and Δ_n denote an arbitrary interval situated in $(\rho_n, 1)$. Then under the hypothesis of Lemma 1 we have*

$$\int_0^{2\pi} \left(\sum_{n=1}^{\infty} | s_{n,k_n}(\rho_n, \theta) |^m \right)^{r/m} d\theta \leqq A \int_0^{2\pi} \left(\sum_{n=1}^{\infty} \frac{1}{\Delta_n} \int_{\Delta_n} | f_n(\rho, \theta) |^m d\rho \right)^{r/m} d\theta.$$

The sums $s_{n,\nu}$ may be replaced by the conjugate ones.

This is immediate from Lemma 2.

PROOF OF (2.2). We have

$$s_n(e^{i\theta}) - \tau_n(e^{i\theta}) = -is_n'(e^{i\theta})/(n+1).$$

By Abel's transformation

$$s_n'(e^{i\theta}) = \rho^{-n} s_n'(\rho e^{i\theta}) - (1-\rho)\sum_{\nu=0}^{n-1} \rho^{-\nu-1} s_\nu'(\rho e^{i\theta}).$$

Therefore

$$|s_n'(e^{i\theta})|^q \leqq A\left\{\rho^{-qn}|s_n'(\rho e^{i\theta})|^q + (1-\rho)^q\left(\sum_{\nu=0}^{n-1}\rho^{-\nu-1}|s_\nu'(\rho e^{i\theta})|\right)^q\right\}$$

$$\leqq B\left\{\rho^{-qn}|s_n'(\rho e^{i\theta})|^q + \frac{1-\rho}{\rho^{(q-1)n}}\sum_{\nu=0}^{n-1}\rho^{-\nu-1}|s_\nu'(\rho e^{i\theta})|^q\right\}.$$

Let $\rho_n = 1-1/(n+1)$ and $\Delta_n = (\rho_n, \rho_{n+1})$ and let I be the left-hand side integral of (2.2). Then we have

$$I \leqq C\int_0^{2\pi}\left(\sum_{n=1}^{\infty}\frac{|s_n'(\rho_n e^{i\theta})|^q}{n^{q+1}\rho_n^{qn}} + \sum_{n=1}^{\infty}\frac{1-\rho_n}{n^{q+1}\rho_n^{(q-1)n}}\sum_{\nu=0}^{n-1}\rho_n^{-\nu-1}|s_\nu'(\rho_n e^{i\theta})|^q\right)^{r/q} d\theta$$

$$\leqq D\int_0^{2\pi}\left(\sum_{n=1}^{\infty}\frac{|s_n'(\rho_n e^{i\theta})|^q}{n^{q+1}\Delta_n}\int_{\Delta_n}|\phi'(\rho e^{i\theta})|^q d\rho + \sum_{n=1}^{\infty}\frac{1}{n^{q+2}}\sum_{\nu=0}^{n-1}\frac{1}{\Delta_n}\int_{\Delta_n}|\phi'(\rho e^{i\theta})|^q d\rho\right)^{r/q} d\theta$$

$$\leqq E\int_0^{2\pi}\left(\sum_{n=1}^{\infty}\frac{n^2\cdot n^{q-1}}{n^{q+1}}\int_{\Delta_n}(1-\rho)^{q-1}|\phi'(\rho e^{i\theta})|^q d\rho\right)^{r/q} d\theta$$

$$\leqq F\int_0^{2\pi}\left(\int_0^1(1-\rho)^{q-1}|\phi'(\rho e^{i\theta})|^q d\rho\right)^{r/q} d\theta$$

$$\leqq G\int_0^{2\pi}|\phi(e^{i\theta})|^r d\theta$$

from (1.2). Thus we obtain (2.2).

PROOF OF (2.3). We may write

$$|\phi'(\rho e^{i\theta})| = \left|\sum_{\nu=1}^{\infty}\nu c_\nu \rho^{\nu-1}e^{i\nu\theta}\right| = (1-\rho)\left|\sum_{\nu=1}^{\infty}s_\nu'\rho^{\nu-1}\right|.$$

If we put $\rho_n = 1-1/n$, then we have

$$\int_0^1 (1-\rho)^{p-1}\,|\,\phi'\,|^p d\rho = \sum_{n=1}^{\infty}\int_{\rho_n}^{\rho_{n+1}} (1-\rho)^{p-1}\,|\,\phi'\,|^p d\rho$$

$$\leqq \sum_{n=1}^{\infty}\int_{\rho_n}^{\rho_{n+1}} (1-\rho)^{p-1}\left[(1-\rho)\sum_{\nu=1}^{\infty}|\,s_\nu'\,|\,\rho^{\nu-1}\right]^p d\rho$$

$$\leqq \sum_{n=1}^{\infty}\frac{1}{n^{p+1}}\left[(1-\rho_n)\sum_{\nu=1}^{\infty}|\,s_\nu'\,|\,\rho_{n+1}^{\nu-1}\right]^p$$

$$\leqq A\sum_{n=1}^{\infty}\frac{1}{n^{2p+1}}\left(\sum_{\nu=1}^{\infty}|\,s_\nu'\,|\,\rho_{n+1}^{\nu-1}\right)^p$$

$$\leqq B\sum_{n=1}^{\infty}\frac{1}{n^{2p+1}}\left(\sum_{\nu=1}^{n}|\,s_\nu'\,|\,\rho_{n+1}^{\nu-1}\right)^p + B\sum_{n=1}^{\infty}\frac{1}{n^{2p+1}}\left(\sum_{\nu=n+1}^{\infty}|\,s_\nu'\,|\,\rho_{n+1}^{\nu-1}\right)^p$$

$$= P + Q$$

say. We have

$$P \leqq \sum_{n=1}^{\infty}\frac{1}{n^{2p+1}}\left(\sum_{\nu=1}^{n}|\,s_\nu'\,|^p\right)\left(\sum_{\nu=1}^{n}\rho_{n+1}^{p'(\nu-1)}\right)^{p/p'}$$

$$\leqq C\sum_{n=1}^{\infty}\frac{1}{n^{2p+1}}\left(\sum_{\nu=1}^{n}|\,s_\nu'\,|^p\right)\cdot n^{p/p'}$$

$$\leqq C\sum_{n=1}^{\infty}\frac{1}{n^{p+2}}\sum_{\nu=1}^{n}|\,s_\nu'\,|^p \leqq C\sum_{\nu=1}^{\infty}\frac{|\,s_\nu'\,|^p}{\nu^{p+1}}$$

where $1/p+1/p'=1$, and

$$Q \leqq \sum_{n=1}^{\infty}\frac{1}{n^{2p+1}}\left(\sum_{\nu=n+1}^{\infty}|\,s_\nu'\,|\,\rho_{n+1}^{\nu-1}\right)^p$$

$$\leqq D\sum_{n=1}^{\infty}\frac{1}{n^{2p+1}}\left(\sum_{\nu=n+1}^{\infty}\frac{|\,s_\nu'\,|^p}{\nu^{p+2}}\right)\left(\sum_{\nu=n+1}^{\infty}\nu^{(1+2/p)p'}\,\rho_{n+1}^{p'(\nu-1)}\right)^{p/p'}$$

$$\leqq D\sum_{n=1}^{\infty}\frac{1}{n^{2p+1}}\left(\sum_{\nu=n+1}^{\infty}\frac{|\,s_\nu'\,|^p}{\nu^{p+2}}\right)\left(\sum_{\nu=n+1}^{\infty}\nu^{2p'-2}\,\rho_{n+1}^{p'(\nu-1)}\right)^{p/p'}$$

$$\leqq D\sum_{n=1}^{\infty}\sum_{\nu=n+1}^{\infty}|\,s_\nu'\,|^p/\nu^{p+2} \leqq E\sum_{\nu=1}^{\infty}|\,s_\nu'\,|^p/\nu^{p+1}.$$

Since $|s_\nu'|/(\nu+1)=|s_\nu-\tau_\nu|$, we get (2.2) by (1.2).

3. THEOREM 3. *Let $f(\theta)$ be a function belonging to L^r $(r>1)$, and $\{p_n\}$ and $\{d_n\}$ be arbitrary increasing sequences of positive integers such that*

$$p_n/d_n = O(p_{n+1}-p_n).$$

Then we have

$$(3.1)\qquad \int_0^{2\pi}\left(\sum_{\nu=1}^{\infty}|s_{p_\nu}-\tau_{p_\nu}|^q/d_\nu\right)^{r/q}d\theta \leqq A\int_0^{2\pi}|f(\theta)|^r d\theta \quad (q\geqq 2).$$

Especially if

$$p_n/n = O(p_{n+1}-p_n),$$

then

$$(3.2)\qquad \sum_{\nu=1}^{\infty}|s_{p_\nu}-f|^m/\nu < \infty \qquad (m\geqq 1)$$

for almost all θ.

The proof of (3.1) runs similarly to that of (2.2). In this case we have to replace Δ_n and ρ_n by

$$\Delta_n = \begin{cases}(1-1/p_n,\ 1-1/p_{n+1}), & \text{if } p_{n+1}<2p_n,\\ (1-1/p_n,\ 1-1/2p_n), & \text{if } p_{n+1}\geqq 2p_n,\end{cases}$$

and $\rho_n=1-1/(p_n+1)$ respectively.

(3.2) is obvious from (3.1) by Kronecker's theorem.

REMARK. The formula

$$1/d(x)\sim cp'(x)/p(x)$$

gives the relation between the strong summability factor $d(x)$ and the lacunary sequence $p(x)$. If we put $d(x)=1$, then $p(x)=2^x$, which is nothing but the Kolmogoroff theorem [2]. If we put $d(x)=x$, $x^{1/2}$, and so on we get $p(x)=x^e$, $2^{ex^{1/2}}$, and so on, respectively.

We conclude with the following theorem which is proved by an argument due to Zygmund [7, 8, 9].

THEOREM 4. *If* $f(\theta)$ *belongs to* L^r $(r>1)$, *then for almost all* θ, *the sequence* $\{1, 2, \cdots\}$ *can be divided into two complementary subsequences* $\{n_k\}$ *and* $\{m_k\}$, *depending in general on* θ, *and such that*

$$s_{p_{n_k}}(\theta) \text{ tends to } f(\theta),$$

the series $\sum 1/d_{m_k}$ *converges,*

where $d_n\uparrow$, $\sum 1/d_n=\infty$, *and*

$$p_n/d_n = O(p_{n+1}-p_n).$$

BIBLIOGRAPHY

1. R. P. Boas, Jr., and S. Bochner, *On a theorem of M. Riesz for Fourier series*, J. London Math. Soc. vol. 14 (1939) pp. 62–73.

2. A. Kolmogoroff, *Une contribution à l'étude de la convergence des séries de Fourier*, Fund. Math. vol. 5 (1924) pp. 96–97.

3. J. E. Littlewood and R. E. A. C. Paley, *Theorems on Fourier series and power series*, J. London Math. Soc. vol. 6 (1931) pp. 230–233.

4. ———, *Theorems on Fourier series and power series*, (2), Proc. London Math. Soc. (2) vol. 42 (1937) pp. 52–89.

5. J. Marcinkiewicz and A. Zygmund, *A theorem of Lusin*, Duke Math. J. vol. 4 (1938) pp. 473–485.

6. G. Sunouchi, *On the strong summability of Fourier series*, Proc. Imp. Acad. Tokyo vol. 19 (1943) pp. 420–423.

7. Z. Zalcwasser, *Sur la sommabilité des séries de Fourier*, Studia Mathematica vol. 6 (1936) pp. 82–88.

8. A. Zygmund, *Sur la caractère de divergence des séries orthogonales*, Mathematica vol. 9 (1935) pp. 86–88.

9. ———, *Proof of a theorem of Paley*, Proc. Cambridge Philos. Soc. vol. 34 (1938) pp. 125–133.

10. ———, *On the convergence and summability of power series on the circle of convergence* (1), Fund. Math. vol. 30 (1938) pp. 170–196.

TOHOKU IMPERIAL UNIVERSITY

ON THEOREMS OF M. RIESZ AND ZYGMUND

A. P. CALDERÓN

Several proofs have been given of the results of M. Riesz and Zygmund.

(a) *The conjugate of the Fourier series of a function $f(x)$ of L^p, $p>1$, is Fourier series of a function $\bar{f}(x)$ of the same class, and*

$$\int_0^{2\pi} |\bar{f}(x)|^p dx \leqq A_p \int_0^{2\pi} |f(x)|^p dx$$

holds, A_p is a constant depending only on p.

(b) *If the function $|f(x)| \log^+ |f(x)|$ is integrable, the conjugate of the Fourier series of $f(x)$ is the Fourier series of a function $\bar{f}(x)$ of the class L. Moreover, there exist two constants A and B such that*

$$\int_0^{2\pi} |\bar{f}(x)|\, dx \leqq A \int_0^{2\pi} |f(x)| \log^+ |f(x)|\, dx + B.$$

In view of the importance of these theorems it may be of interest to give another proof of them based on a different idea. Actually it is

Received by the editors May 14, 1949.

easy to see (and very well known) that they are equivalent to the following proposition about analytic functions.

Let $F(z)=u(z)+iv(z)$, $u(z)>0$, $v(0)=0$ be a function regular inside the circle $|z|<1$. Then the inequalities

$$(1)\qquad \int_0^{2\pi} |v(re^{i\varphi})|^p d\varphi \leqq A_p \int_0^{2\pi} |u(re^{i\varphi})|^p d\varphi; \qquad 1<p\leqq 2;\ r<1,$$

$$(2)\qquad \int_0^{2\pi} |v(re^{i\varphi})|\, d\varphi \leqq C\int_0^{2\pi} |u(re^{i\varphi})|\, d\varphi + D\int_0^{2\pi} u\log u\, d\varphi - 2\pi D u(0)\log u(0)$$

hold, C and D being positive constants and A_p the same as above.

Let us prove (1) first. The argument is based on the inequality

$$(3)\qquad |\sin\phi|^p \leqq A_p|\cos\phi|^p - B_p\cos p\phi$$

valid for $-\pi/2\leqq\phi\leqq\pi/2$, $1<p\leqq 2$, A_p and B_p being two positive constants depending on p.

The proof of it is simple. Since $\cos p\phi<0$ for $\phi=\pm\pi/2$, taking B_p large enough we get $-B_p\cos p\phi>1$ in a small neighborhood of $\phi=\pm\pi/2$. Now $|\cos\phi|^p$ is greater than a positive constant in every closed interval interior to $(-\pi/2,\pi/2)$, so that choosing A_p large enough we shall have $A_p|\cos\phi|^p-B_p\cos p\phi>1$ in $(-\pi/2,\pi/2)$ and a fortiori the inequality (3).

To show (1) now, let us set

$$F(z)=Re^{i\phi};\qquad u=R\cos\phi;\qquad v=R\sin\phi.$$

Owing to the fact that $u>0$, we may choose $-\pi/2<\phi<\pi/2$. For the same reason, $F(z)\neq 0$ and so $F(z)^p$ is regular in $|z|<1$. Integrating now along the circle $|z|=r$ we get

$$\Re\left[\frac{1}{2\pi i}\int_C \frac{F(z)^p}{z}\,dz\right]=\frac{1}{2\pi}\int_0^{2\pi} R^p\cos p\phi\, d\varphi=\Re[F(0)^p]=u(0)^p>0.$$

Multiplying now (3) by R^p and integrating, we obtain

$$\int_0^{2\pi} |v(re^{i\varphi})|^p d\varphi \leqq A_p\int_0^{2\pi} |u(re^{i\varphi})|^p d\varphi - B_p\int_0^{2\pi} R^p\cos p\phi\, d\varphi,$$

but since the last term is positive, we may drop it, and the desired result follows. A similar argument works in proving (2). Here instead of (3) we use

$$(4) \qquad |\sin\phi| \leqq C\cos\phi + D(\cos\phi \log\cos\phi + \phi\sin\phi),$$

which can be proved in exactly the same way.

Again, since $F(z)$ does not vanish in $|z|<1$, $F \log F$ is regular there, and integrating along $|z|=r$ we get

$$\Re\left[\frac{1}{2\pi i}\int_C F(z)\log F(z)\,\frac{dz}{z}\right] = \frac{1}{2\pi}\int_0^{2\pi} \Re[F\log F]dx = u(0)\log u(0),$$

and replacing

$$\begin{aligned}\Re[F\log F] &= \Re[R(\cos\phi + i\sin\phi)(\log R\cos\phi - \log\cos\phi + i\phi)] \\ &= R\cos\phi\log R\cos\phi - R(\cos\phi\log\cos\phi + \phi\sin\phi),\end{aligned}$$

we obtain

$$\frac{1}{2\pi}\int_0^{2\pi} R(\cos\phi\log\cos\phi + \phi\sin\phi)dx = \frac{1}{2\pi}\int_0^{2\pi} u\log u dx - u(0)\log u(0).$$

Finally, multiplying (4) by R and integrating, we obtain

$$\int_0^{2\pi} |v(re^{ix})|\,dx \leqq C\int_0^{2\pi} |u(re^{ix})|\,dx + D\int_0^{2\pi} u\log u dx - 2\pi D u(0)\log u(0),$$

as stated in (2).

Bibliography

1. M. Riesz, *Sur les fonctions conjuguées*, Math. Zeit. vol. 27 (1927) pp. 218–244.
2. A. Zygmund, *Sur les fonctions conjuguées*, Fund. Math. vol. 13 (1929) pp. 284–303.
3. ———, *Trigonometrical series*, Warsaw, 1935.

University of Chicago

SUMMABILITY AND ANALYTIC CONTINUATION

V. F. COWLING[1]

1. **Introduction.** In this paper we describe a new family of Toeplitz summability methods, and we study the regions in which these methods sum a Taylor series to the analytic continuation of the function which it represents.

Let $A=(a_{nm})$ and $x=\{s_m\}$ $(n, m=0, 1, \cdots)$ be a matrix and a sequence of complex numbers, respectively. We write formally

$$t_n \equiv A_n(x) \equiv \sum_{m=0}^{\infty} a_{nm}s_m, \tag{1}$$

and say that the sequence x (and the corresponding series $\sum_{m=0}^{\infty}(s_m-s_{m-1})$, with $s_{-1}=0$) is summable A to the sum t if each of the series in (1) converges and $\lim t_n$ exists and equals t. We say that the method A is regular provided it sums every convergent sequence to its limit. The method A is regular if and only if

$$\sum_{m=0}^{\infty} |a_{nm}| \leqq K \qquad (n = 0, 1, \cdots), \tag{2}$$

$$\lim_{n\to\infty} a_{nm} = 0 \qquad (m = 0, 1, \cdots), \tag{3}$$

$$\lim_{n\to\infty} \sum_{m=0}^{\infty} a_{nm} = 1, \tag{4}$$

where K is a constant independent of n (cf. Toeplitz[2] [2]).

2. **The methods F^r.** For each constant r $(r\neq 1)$ the element of the matrix $A=F^r$ shall be defined by the equations

$$\begin{aligned} a_{nm}(r) &= 0 && (m < n), \\ a_{nm}(r) &= (1-r)^{n+1}C_{m,n}r^{m-n} && (m \geqq n). \end{aligned} \tag{5}$$

We note first that, for $n=0, 1, \cdots$ and $|r|<1$,

$$\sum_{m=0}^{\infty} a_{nm}(r) = (1-r)^{n+1}\sum_{m=n}^{\infty} C_{m,n}r^{m-n} = 1$$

Presented to the Society, December 30, 1948; received by the editors June 2, 1949.

[1] The author wishes to acknowledge the constructive criticism of Professor G. Piranian on this paper.

[2] Numbers in brackets refer to the bibliography at the end of the paper.

and

$$\sum_{m=0}^{\infty} |\, a_{nm}(r)\,| = \left(\frac{1-r}{1-|r|}\right)^{n+1}.$$

If on the other hand $|r| \geqq 1$ $(r \neq 1)$, then $\sum_{m=0}^{\infty} |a_{nm}(r)| = \infty$ $(n = 0, 1, \cdots)$. Therefore the method F^r is regular if and only if r is real and satisfies the condition $0 \leqq r < 1$.

THEOREM[3] 2.1. *If AB denotes the matrix product of A and B, $F^{r_1}F^{r_2} = F^{r_2}F^{r_1} = F^r$, where $r = r_1 + r_2 - r_1r_2$.*

Let $F^{r_1} = (a_{nm})$, $F^{r_2} = (b_{nm})$, $F^{r_1}F^{r_2} = (c_{nm})$ and $F^{r_2}F^{r_1} = (\gamma_{nm})$. Then $c_{nm} = 0$ *when* $n > m$. *When* $n \leqq m$,

$$\begin{aligned}
c_{nm} &= \sum_{k=0}^{\infty} a_{nk}b_{km} \\
&= (1-r_1)^{n+1} r_1^{-n} r_2^{m} \sum_{k=n}^{m} r_1^{k} r_2^{-k} (1-r_2)^{k+1} C_{k,m} C_{n,k} \\
&= (1-r_1)^{n+1} r_1^{-n} r_2^{m} \left(\frac{r_1}{r_2}\right)^{n} (1-r_2)^{n+1} C_{m,n} \sum_{k=0}^{m-n} C_{m-n,k} \left(\frac{r_1}{r_2} - r_1\right)^{k} \\
&= (1-r_1)^{n+1}(1-r_2)^{n+1} C_{m,n} (r_1 + r_2 - r_1r_2)^{m-n} \\
&= [1 - (r_1 + r_2 - r_1r_2)]^{n+1} C_{m,n} (r_1 + r_2 - r_1r_2)^{m-n}.
\end{aligned}$$

Since the last expression is symmetric in r_1 and r_2, $\gamma_{nm} = c_{nm}$, the proof is complete. We note the corollary that, for $r \neq 1$, the matrix F^r has the inverse F^ρ, where $\rho = -r/(1-r)$.

THEOREM 2.2. *If r_1 and r_2 are real constants $(0 \leqq r_1 < r_2 < 1)$, every bounded sequence summed by F^{r_1} is summed by F^{r_2}, and to the same sum.*

Let $r = (r_2 - r_1)/(1 - r_1)$; then the method F^r is regular, and $F^{r_2} = F^rF^{r_1}$, by Theorem 2.1. The present theorem now follows from a remark by Agnew [1] (cf. p. 328).

3. **Summability of Taylor series.** Henceforth, $\{u_m\}$ and $\{s_m\}$ shall denote sequences of complex numbers related by the equations

$$s_m = u_0 + u_1 + \cdots + u_m \qquad (m = 0, 1, \cdots).$$

LEMMA. *If*

[3] The author is indebted to A. Wilansky for helpful discussion regarding this result.

(i) *the series* $\sum u_m z^m$ *has a positive radius of convergence* R;
(ii) $|r| < R$;
(iii) $t_n = (1-r)^n \sum_{m=n}^{\infty} C_{m,n} U_m r^{m-n}$ $(n = 0, 1, \cdots)$;
(iv) $T_n = t_0 + t_1 + \cdots + t_n$ $(n = 0, 1, \cdots)$;
then

$$T_n = (1-r)^{n+1} \sum_{m=n}^{\infty} C_{m,n} s_m r^{m-n} \qquad (n = 0, 1, \cdots). \tag{6}$$

Equation (6) asserts that a certain sum of n Taylor series is equal to another Taylor series. All the series in question converge because, for any fixed positive integer n, the respective radii of convergence of the four series

$$\sum_{m=0}^{\infty} u_m z^m, \qquad \sum_{m=0}^{\infty} s_m z^m, \qquad \sum_{m=0}^{\infty} C_{m,n} u_m z^m, \qquad \sum_{m=0}^{\infty} C_{m,n} s_m z^m$$

are equal. The validity of the equation will now be established by induction.

For $n = 0$, the equation (6) reduces to the identity

$$(1-r) \sum_{m=0}^{\infty} s_m r^m = \sum_{m=0}^{\infty} u_m r^m.$$

But if the equation holds for $n = k$, that is, if

$$T_k = \frac{(1-r)^{k+1}}{r^k} \sum_{m=k}^{\infty} C_{m,k} s_m r^m,$$

then

$$\begin{aligned}
T_{k+1} &= T_k + t_{k+1} \\
&= \left(\frac{1-r}{r}\right)^{k+1} \left\{ \sum_{m=k}^{\infty} C_{m,k} s_m r^{m+1} + \sum_{m=k+1}^{\infty} C_{m,k+1}(s_m - s_{m-1}) r^m \right\} \\
&= \left(\frac{1-r}{r}\right)^{k+1} \left\{ \sum_{m=k+1}^{\infty} [C_{m,k} r + C_{m,k+1} - C_{m+1,k+1} r] s_m r^m \right. \\
&\qquad \left. + [C_{k,k} - C_{k+1,k+1}] s_k r^{k+1} \right\} \\
&= \left(\frac{1-r}{r}\right)^{k+1} \sum_{m=k+1}^{\infty} \{C_{m,k+1} - [C_{m+1,k+1} - C_{m,k}] r\} s_m r^m \\
&= \frac{(1-r)^{k+2}}{r^{k+1}} \sum_{m=k+1}^{\infty} C_{m,k+1} s_m r^m.
\end{aligned}$$

In other words, equation (6) also holds for $n-k+1$. This proves the lemma. From the lemma it follows that if $0<|r|<R$, where R is the radius of convergence of the series $\sum u_m z^m$, the relation

$$\sum_{n=0}^{\infty}\left(\frac{1-r}{r}\right)^n \sum_{m=n}^{\infty} C_{m,n}u_m r^m = \lim_{n\to\infty}\frac{(1-r)^{n+1}}{r^n}\sum_{m=n}^{\infty} C_{m,n}s_m r^m$$

holds in the sense that the existence of either member implies the existence, with the same value, of the other member. The following result is now immediate:

THEOREM 3.1. *If the series $\sum_{m=0}^{\infty} u_m z^m$ has a positive radius of convergence R, the series $\sum u_m$ is summable F^r to the sum L for any constant r $(0<|r|<R)$ for which the series*

$$\sum_{m=0}^{\infty}\left(\frac{1-r}{r}\right)^n \sum_{m=n}^{\infty} C_{m,n}u_m r^m$$

converges to L.

4. **Analytic continuation by means of the methods F^r.** Let D be a simply connected region in the complex plane, C a simple closed Jordan curve lying in D and bounding the finite region D'. Corresponding to each complex number r $(r\neq 1)$ we define an open set $R(r, C)$ as follows: $R(r, C)$ is the set of all points in D' for which $|z-rz|<|t-rz|$ whenever t lies on C. A set B shall be said to be of type $R^*(r, D)$ if B is closed and if C can be chosen in such a way that B is a subset of $R(r, C)$. In the following theorem, $f^{(n)}(rz)$ denotes the nth derivative with respect to w of $f(w)$ at the point $w=rz$.

THEOREM 4.1. *If $f(z)$ is uniform and regular in the region D, r is a complex constant $(r\neq 1)$, and B is a set of type $R^*(r, D)$, then the series*

$$\sum_{n=0}^{\infty} f^{(n)}(rz)(1-r)^n z^n/n! \tag{7}$$

converges to $f(z)$ uniformly in B.

If z lies in B,

$$f(z)=\frac{1}{2\pi i}\int_C \frac{f(t)dt}{t-z}=\frac{1}{2\pi i}\int_C \frac{f(t)}{t-rz}\,\frac{1}{1-(z-rz)/(t-rz)}\,dt$$

$$=\frac{1}{2\pi i}\int_C \frac{f(t)}{t-rz}\sum_{n=0}^{\infty}\left(\frac{z-rz}{t-rz}\right)^n dt,$$

where C is a curve in D such that B is a subset of $R(r, C)$. The series

in the last integrand converges uniformly with respect to t on C and z in B, because, for such values of t and z,

$$|(z - rz)/(t - rz)| < K < 1,$$

where the constant K depends only on B and C. Therefore

$$f(z) = \sum_{n=0}^{\infty} (z - rz)^n \frac{1}{2\pi i} \int_C \frac{f(t)dt}{(t - rz)^{n+1}}.$$

But the point rz is inside of the curve C; for otherwise the line segment joining the points z and rz would meet the curve C at some point t', the inequality $|(z-rz)/(t'-rz)| < 1$ would not be satisfied, and the point z would not lie in B. Therefore the last integral has the value $2\pi i f^{(n)}(rz)/n'$, and the theorem is proved.

The hypothesis that the region D is simply connected will now be removed. Let D be a connected open set, E its boundary, and r a complex constant ($r \neq 1$). By $Q(r, D)$ we shall denote the set of all points in D for which the inequality $|z-rz| < |t-rz|$ is satisfied whenever t lies in E.

THEOREM 4.2. *Let the function $f(z)$ be regular and uniform in the bounded region D, and let r be a complex constant ($r \neq 1$). Then the series (7) converges to $f(z)$ absolutely in $Q(r, D)$, and the absolute convergence is uniform in every closed subset of $Q(r, D)$.*

Let B be a closed subset of $Q(r, D)$, and z_0 a point in B. It is to be shown that the series

$$\sum_{n=0}^{\infty} f^{(n)}(rz_0)(1 - r)^n z_0^n/n! \tag{8}$$

converges to $f(z_0)$, and that the absolute convergence of the series (9) is uniform with respect to z_0 in B.

We observe that (8) is the Taylor series of $f(z_0)$ about the point $z = rz_0$.[4] Since B is a closed subset of $Q(r, D)$, there exists a positive number ϵ such that $|(z_0 - rz_0)/(t - rz_0)| < 1 - 2\epsilon$ when z_0 is in B and t is on E. The Cauchy estimate for the coefficients of our Taylor series gives the result

$$|f^{(n)}(rz_0)/n!| < M(rz_0)/[(1 - \epsilon)\delta_0]^n,$$

where δ_0 is the distance from rz_0 to the set E and $M(rz_0)$ denotes the maximum modulus of $f(z)$ on the circle $|z - rz_0| = (1-\epsilon)\delta_0$. As the

[4] The author is indebted to A. M. Gleason who pointed out this fact in an oral communication.

point z_0 ranges over the set B, the points

$$z = rz_0 + e^{i\theta}(1 - \epsilon)\delta_0 \qquad (0 \leqq \theta < 2\pi)$$

range over a subset of D which is bounded away from the set E, and therefore the quantities $M(rz_0)$ have a common finite upper bound M. It follows that

$$\left| f^{(n)}(rz_0)(1 - r)^n \frac{z^n}{n!} \right| < \frac{[(1 - 2\epsilon)\delta_0]^n M}{[(1 - \epsilon)\delta_0]^n} < (1 - \epsilon)^n M,$$

and the theorem is proved.

THEOREM 4.3. *Let D be a connected open set containing the origin, $f(z)$ a uniform function regular in D; let the series $\sum a_m z^m$ converge to $f(z)$ in a circle of radius ρ; and let r be a complex constant $(r \neq 1)$. Then, in every closed set which is contained in the intersection of the set $Q(r, D)$ and the circle $|z| < \rho/|r|$, the series $\sum a_m z^m$ is uniformly absolutely summable F^r to $f(z)$.*

If $r = 0$, the set $Q(r, D)$ is contained in the circle $|z| < \rho$, and the result is trivial. To prove the theorem for $r \neq 0$, we observe first that the series

$$\sum_{m=n}^{\infty} a_m m(m - 1) \cdots (m - n + 1)(rz)^{m-n}$$

converges to $f^{(n)}(rz)$ in the region $|z| < \rho/|r|$. It follows that the series (7) can be written in the form

$$\sum_{n=0}^{\infty} \sum_{m=n}^{\infty} a_m C_{m,n}(1 - r)^n z^m r^{m-n} = \sum_{n=0}^{\infty} \left(\frac{1 - r}{r} \right)^n \sum_{m=n}^{\infty} C_{m,n} r^m a_m z^m,$$

and upon application of Theorem 3.1 (with $u_m = a_m z^m$) the present theorem becomes a corollary of Theorem 4.2. We note that we have established analytic continuation of the function represented by the series $\sum a_m z^m$ by means of summability methods which need not even be regular.

5. **Summability of a special series.** We now subject the methods F^r to the customary test of applying them to the series $\sum z^n$. Here, the sole singularity of the function $f(z)$ is the point $z = 1$, and the boundary of the region D can therefore be taken to consist of the point $z = 1$ together with the circle $|z| = K$, where K is arbitrarily large. We shall restrict our considerations to the case where F^r is regular, that is, where r is real and $0 \leqq r < 1$. If z is any point in the plane, the

inequality $|z-rz| < |t-rz|$ is certainly satisfied if t is sufficiently large. It remains to examine the inequality for the case where $t=1$. Here we have, with $z=x+iy$,

$$(x^2+y^2)(1-r)^2 < (1-rx)^2 + r^2y^2,$$
$$(1-2r)(x^2+y^2) < 1-2rx.$$

Theorem 4.3 now gives the following result: If $0<r<1/2$, F^r sums the series $\sum z^n$ to the function $1/(1-z)$ in the intersection of the region $|z|<1/r$ with the interior of the circle having its center at $z=-r/(1-2r)$ and passing through the point $z=1$. If $1/2<r<1$, F^r sums the series in the intersection of the region $|z|<1/r$ with the exterior of the circle having its center at $z=r/(2r-1)$ and passing through the point $z=1$. The method $F^{1/2}$ sums the series in the intersection of the region $|z|<2$ with the half-plane $x<1$.

Finally, we recall that in §2 the relation $F^{r_2} \supset F^{r_1}$ $(0 \leqq r_1 < r_2 < 1)$ was shown to hold in the space of bounded sequences. The relation does not hold in the space of all sequences. For we have now established that the series $\sum(-5/3)^n$ is summable $F^{1/2}$; on the other hand, the transform of this series by F^r does not even exist when $r \geqq 3/5$.

Added in proof: After the present paper had been accepted for publication, an interesting paper by P. Vermes [Amer. J. Math. vol. 71 (1949) pp. 541–562] appeared which overlaps the results of the present paper. In addition R. P. Agnew [Math. Rev. vol. 11, p. 242] in a review of a paper on these methods by W. König-Meyer [Math. Zeit. vol. 52 (1949) pp. 257–304] points out that the first systematic study of these methods is in the thesis of R. Wais [*Das Taylorsche Summierungsverfahren*, Tubingen, 1935].

Bibliography

1. R. P. Agnew, *Comparison of products of methods of summability*, Trans. Amer. Math. Soc. vol. 43 (1938) pp. 327–343.

2. O. Toeplitz, *Über allgemeine lineare Mittelbildungen*, Prace Matematyczno-Fizyczne vol. 22 (1911) pp. 113–119.

Lehigh University

A BESSEL FUNCTION EXPANSION

MORRIS KLINE[1]

1. **Introduction.** Some work in the theory of electromagnetic wave guides led to the following problem, namely, to obtain a useful approximation for the expression

$$J_\nu(kR)N_\nu(kR+ka) - J_\nu(kR+ka)N_\nu(kR), \tag{1}$$

wherein J_ν and N_ν are the Bessel functions of the first and second kind of order ν, and under the conditions that R be large and that $\nu = dR$, where d is any real constant. It is perhaps worth while to record the result, particularly because it is obtained by rather elementary means. The result is a Taylor's expansion of the expression (1) about $R=\infty$ as follows:

$$J_\nu(kR)N_\nu(kR+ka) - J_\nu(kR+ka)N_\nu(kR) \tag{2}$$
$$= \frac{2}{\pi}\left[\frac{a}{R}\,\frac{\sin x}{x} + \frac{a^2}{2R^2}\left\{\frac{\sin x}{\gamma x} - \left(1+\frac{1}{\gamma}\right)\cos x\right\} + \cdots\right]$$

wherein $x^2 = -k^2a^2\gamma$, $\gamma = (d^2/k^2) - 1$ and $\nu = dR$. A similar result is also noted below for the expression:

$$J_\nu'(kR)N_\nu'(kR+ka) - J_\nu'(kR+ka)N_\nu'(kR). \tag{3}$$

Physical interpretation of the two results is given in §3.

It might be remarked that it did not appear possible to obtain any simple result with the usual asymptotic expansions for the individual Bessel functions.

2. **The derivation of the basic expansion.** Let $\alpha = ka$ and $\beta = kR$ in (1) and let us denote it by $G(\alpha, \beta)$. We shall, for the time being, regard the order ν of the Bessel functions J_ν and N_ν as fixed and expand $N_\nu(\beta+\alpha)$ and $J_\nu(\beta+\alpha)$ in powers of α at the point β. We obtain:

$$\begin{aligned} G(\alpha,\beta) &= J_\nu(\beta)\{N_\nu(\beta) + \alpha N_\nu'(\beta) + (\alpha^2/2)\,N_\nu''(\beta) + \cdots\} \\ &\quad - N_\nu(\beta)\{J_\nu(\beta) + \alpha J_\nu'(\beta) + (\alpha^2/2)\,J_\nu''(\beta) + \cdots\}. \end{aligned} \tag{4}$$

The functions $J_\nu(z)$ and $N_\nu(z)$ are analytic except at $z=\infty$ and, in the case of the second one, at $z=0$. The Taylor series representation is therefore valid for all values of α and β chosen subject to these restric-

Received by the editors November 24, 1948 and, in revised form, April 27, 1949.

[1] Work on this paper was made possible by funds provided by the Watson Laboratories, Air Materiel Command, United States Air Force, under a contract with New York University.

tions; the series must converge absolutely; and by an allowable rearrangement of its terms $G(\alpha, \beta)$ may be written as

$$G(\alpha, \beta) = \sum_{n=1}^{\infty} \frac{\alpha^n}{n!} D_{0,n}(\beta) \tag{5}$$

where we have adopted the notation, with respect to any variable z,

$$D_{i,j}(z) = J_\nu^{[i]}(z) N_\nu^{[j]}(z) - J_\nu^{[j]}(z) N_\nu^{[i]}(z), \tag{6}$$

the superscript $[i]$ meaning the ith derivative with respect to z. We shall investigate the expression (5) with a view to ultimately finding those terms which may justifiably be regarded as the *leading* terms in $1/R$ of the expansion (5) under the condition $\nu = dR$.[2]

Let $u(z)$ and $v(z)$ denote any two independent solutions of the Bessel equation of order ν. Then $u'' = -z^{-1}u' + z^{-2}(\nu^2 - z^2)u$, and $v'' = -z^{-1}v' + z^{-2}(\nu^2 - z^2)v$. Substituting these equations into $D_{2,\nu}$ gives immediately that $D_{2,\nu} = -z^{-1}D_{1,\nu} + z^{-2}(\nu^2 - z^2)D_{0,\nu}$. If we set:

$$\rho = z^{-1}, \qquad \gamma = z^{-2}(\nu^2 - z^2) = \nu^2\rho^2 - 1, \tag{7}$$

we may write

$$D_{2,\nu} = -\rho D_{1,\nu} + \gamma D_{0,\nu}. \tag{8}$$

It is immediate, from equation (6), that if we let the symbol δ stand for differentiation with respect to z then:

$$\delta D_{i,j} = D_{i+1,j} + D_{i,j+1}. \tag{9}$$

By use of formulas (9) and (8) we readily establish that

$$D_{0,n} = (2\delta + \rho)D_{0,n-1} + (\gamma - \rho\delta - \delta^2)D_{0,n-2}. \tag{10}$$

We introduce the notation $\Delta_n = (\pi/2)D_{0,n}$, the factor $\pi/2$ being chosen to cancel a $2/\pi$ which will appear later in every $D_{0,n}$. Also we introduce the operators:

$$\sigma = 2\delta + \rho \quad \text{and} \quad \tau = \gamma - \rho\delta - \delta^2. \tag{11}$$

Then relation (10) above becomes the basic recursion formula:

$$\Delta_n = \sigma\Delta_{n-1} + \tau\Delta_{n-2}. \tag{12}$$

The formulas for Δ_n for the first few integers n are well known,[3] and

[2] A special case of expression (5) for the case $\nu = 0$ is used in Schelkunoff, S., *The electromagnetic theory of coaxial transmission lines and cylindrical shields*, Bell System Technical Journal vol. 13 (1934) p. 532 ff. See p. 562 in particular.

[3] For $n = 0, 1, \cdots, 5$, the function $D_{0,n}$ can be found in Watson, G. N., *Introduction to the theory of Bessel functions*, New York, Macmillan, 1945, p. 76. Formula (12) provides a convenient way of deriving as many more as may be needed.

easily derived. The first two are $\Delta_0=0$ and $\Delta_1=\rho$. From the first two, and the recursion relation (12), we see that $\Delta_2=-\rho^2$, since ρ is z^{-1}. In general, the expression Δ_n is a polynomial in ρ and ν of increasing order in ρ and ν. It is also a polynomial in ρ and $\gamma=(\nu^2\rho^2-1)$. The form of these polynomials becomes exceedingly complicated with increasing n.

We are interested in the dependence of $G(\alpha, \beta)$ on $1/R$ as $R\to\infty$ and $\nu=dR$. To obtain this relationship we intend to let $1/\rho=z=kR$, whereupon the combination $\nu\rho$, and therefore γ, will be of zero order in $1/R$. The variable ρ itself will be of first order in $1/R$. Guided by these facts we shall seek out in each Δ_n the terms of total orders one and two in $1/R$, discarding those of higher order.

We shall prove by induction that for all $n\geqq 1$

$$(13)\quad \begin{aligned} &\text{(a)}\ \ \Delta_{2n-1}\sim\rho\gamma^{n-1};\\ &\text{(b)}\ \ \Delta_{2n}\sim -n(2n-1)\rho^2\gamma^{n-1}-2n(n-1)\rho^2\gamma^{n-2}, \end{aligned}$$

where the explicit terms on the right are exact, and where all terms omitted are of the form $P\rho^{s+2}\gamma^m$, $s>0$, P a function of n. There are a finite number of such terms (for each n) and each of them, as remarked, is of at least third order in the ultimate variable $1/R$.

By the definitions of σ and τ it follows that for every $m\geqq 0$,

$$(14)\quad \begin{aligned} &\text{(i)}\ \ \sigma\{\rho\gamma^m\}=-(1+4m)\rho^2\gamma^m-4m\rho^2\gamma^{m-1};\\ &\text{(ii)}\ \ \tau\{\rho\gamma^m\}\sim\rho\gamma^{m+1};\quad \text{(iii)}\ \ \sigma\{\rho^2\gamma^m\}\sim 0;\\ &\text{(iv)}\ \ \tau\{\rho^2\gamma^m\}\sim\rho^2\gamma^{m+1};\\ &\text{(v)}\ \ \sigma\{\rho^{s+2}\gamma^m\}\sim 0\quad\text{and}\quad \tau\{\rho^{s+2}\gamma^m\}\sim 0, \qquad s>0; \end{aligned}$$

the symbol $\sim$ being understood as directly above.

Turning now to formulas (13), note first that they are true in the special case $n=1$, and we must verify that if they hold for a given n then they hold also for $n+1$. By the recursion formula (12), $\Delta_{2n+1}=\sigma\Delta_{2n}+\tau\Delta_{2n-1}$. Assuming that formula (13) holds for n, we obtain $\Delta_{2n+1}=-n(2n-1)\sigma\{\rho^2\gamma^{n-1}\}-2n(n-1)\sigma\{\rho^2\gamma^{n-2}\}+\{\rho\gamma^{n-1}\}$. Therefore, by formulas (14), $\Delta_{2n+1}\sim\rho\gamma^n$. Notice that in the case $n=1$, where the middle term involving γ^{n-2} might seem disquieting, this term is not actually present because of the factor $(n-1)$ in its coefficient. This proves formula (13a) for Δ_{2n+1}.

To prove formula (13b) for Δ_{2n+2} we write formula (12) in the form $\Delta_{2n+2}=\sigma\Delta_{2n+1}+\tau\Delta_{2n}$. Now we substitute for Δ_{2n+1} from the preceding paragraph and for Δ_{2n} from formula (13b). Then, Δ_{2n+2}

$\sim\sigma\{\rho\gamma^n\}+\tau\{-n(2n-1)\rho^2\gamma^{n-1}-2n(n-1)\rho^2\gamma^{n-2}\}$. Finally by formulas (14), $\Delta_{2n+2}\sim-(n+1)(2n+1)\rho^2\gamma^2-2(n+1)n\rho^2\gamma^{n-1}$. This concludes the demonstration of both formulas (13).

We now let $z=\beta$ (or kR) and substitute the approximations to Δ_n given by formulas (13) in the series (5), after replacing $D_{0,n}$ by $(2/\pi)\Delta_n$. The use of approximations (13) in the terms of this series amounts to regarding the series as one in two variables α and β and to a reordering of the terms. We postpone the justification for this step[4] and turn at once to the series.

First we separate $\sum_{n=1}^{\infty}\alpha^n/n!\Delta_n$ into the sum of an "odd" and "even" series, $\sum_{m=1}^{\infty}\alpha^{2m-1}/(2m-1)!\Delta_{2m-1}$ and $\sum_{m=1}^{\infty}\alpha^{2m}/(2m)!\Delta_{2m}$. Then we substitute for each of the Δ's the value from (13). Hence

$$\sum_{n=1}^{\infty}\frac{\alpha^n}{n!}\Delta_n=\frac{1}{\beta}\sum_{m=1}^{\infty}\frac{\alpha^{2m-1}}{(2m-1)!}\gamma^{m-1}-\frac{1}{\beta^2}\sum_{m=1}^{\infty}\frac{\alpha^{2m}}{(2m)!}\left[m(2m-1)\gamma^{m-1}+2m(m-1)\gamma^{m-2}\right]+\text{remainder.}$$

Each term of this remainder is of the form $(1/\beta^s)f_s(\alpha,\gamma)$, $s\geqq 3$, with f_s an appropriate series in α whose coefficients are polynomials in γ.

We introduce $x^2=-\alpha^2\gamma$; as we shall see later, x is real and positive in applications. By simple manipulations using, in the last series, the relation $n-1=(2n-1)/2-1/2$, we may express this as

$$\sum_{n=1}^{\infty}\frac{\alpha^n}{n!}\Delta_n=\frac{\alpha}{x\beta}\sum_{m=1}^{\infty}\frac{(-1)^{m-1}x^{2m-1}}{(2m-1)!}-\frac{\alpha^2}{2\beta^2}\sum_{m=1}^{\infty}\frac{(-1)^{m-1}x^{2m-2}}{(2m-2)!}+\frac{1}{2}\frac{\alpha^2}{\gamma x\beta^2}\sum_{m=1}^{\infty}\frac{(-1)^{m-1}x^{2m-1}}{(2m-1)!}-\frac{\alpha^2}{2\gamma\beta^2}\sum_{m=1}^{\infty}\frac{(-1)^{m-1}x^{2m-2}}{(2m-2)!}+\text{remainder.}$$

Substituting the trigonometric functions appropriate to the summations gives

$$G(\alpha,\beta)=\frac{2}{\pi}\sum_{1}^{\infty}\frac{\alpha^n}{n!}\Delta_n=\frac{2\alpha}{\beta\pi}\frac{\sin x}{x}+\frac{2\alpha^2}{\pi\beta^2}\left[\frac{\sin x}{\gamma x}-\left(1+\frac{1}{\gamma}\right)\cos x\right]+\text{remainder.}\tag{15}$$

[4] The justification is given in the appendix. The appendix also proves that the expression (2) is a Taylor's expansion about $R=\infty$.

This is the result given in formula (2).

The notation in formula (6) and the formulas (12) and (13) are adequate to treat expression (3). If we again let $\alpha=ka$ and $\beta=kR$ and denote the expression by $\Gamma(\alpha, \beta)$, we obtain

$$\Gamma(\alpha, \beta) = -\frac{2}{\pi}\frac{\alpha\gamma}{\beta}\frac{\sin x}{x} + \frac{\alpha^2}{\pi\beta^2}\left[(\gamma+1)\left(\frac{\sin x}{x}+\cos x\right)+\frac{\gamma \sin x}{x}\right] + \text{remainder}. \tag{16}$$

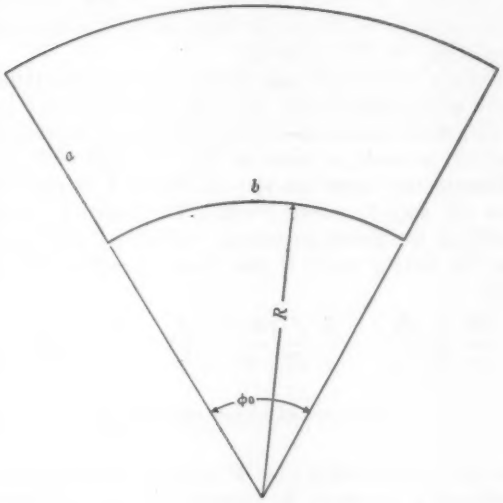


FIG. 1

3. **Physical interpretation of the result.** Given a sector of a coaxial guide, the cross-section of which is indicated in Fig. 1, the longitudinal component of the electrical field supported by such a guide is given by the expression[5]

$$E_z = [KN_\nu(kR)J_\nu(kr) - KJ_\nu(kR)N_\nu(kr)] \sin \nu\phi e^{i\omega t+hz} \tag{17}$$

where r, ϕ, and z are the cylindrical coordinates employed, $\nu=Rm\pi/b$, and $b=R\phi_0$. The quantity K is determined by the amplitude of the

[5] S. Ramo and J. R. Whinnery, *Fields and waves in modern radio*, New York, Wiley, 1944, pp. 325–326 and pp. 333–335.

field excited in the guide. k is determined by setting the left side of equation (2) equal to 0. This equation yields an infinite number of values of k for any one value of ν, each value of k thus determining one mode of the guide.

It is, incidentally, true that setting the right-hand side of equation (2) equal to zero gives an implicit equation for k from which it is possible to determine k to two terms. However, this value of k may also be obtained in other ways. We shall merely note for later reference that if we do set this right-hand expression equal to 0, multiply by R, and let R become infinite, we obtain $\sin x=0$. Since $x=(-\alpha^2\gamma)^{1/2}=a(k^2-d^2)^{1/2}$, we must have $a(k^2-d^2)^{1/2}=n\pi$, $n=0$, ± 1, ± 2. However $d=m\pi/b$ and so $k\big|_{R=\infty}=(n^2\pi^2/a^2+m^2\pi^2/b^2)^{1/2}$, which is the formula for the cutoff frequency of a rectangular waveguide supporting the $TM_{m,n}$ mode.

If one holds a and b fixed and allows R to become infinite while $R\phi_0=b$, one sees geometrically that the coaxial sector becomes a rectangle with dimensions a and b. The physical interpretation of the basic result (2) is made in terms of this rectangle as follows. As R becomes infinite the factor $\sin \nu\phi=\sin Rm\pi\phi/b$ in expression (17) approaches $\sin m\pi y/b$, where $y=R\phi$. The limiting value of the factor involving the Bessel functions is given by the result of this paper; for, by letting $t=r-R$ this factor becomes, by virtue of formula (2),

$$\frac{2K}{\pi}\,\frac{t}{R}\,\frac{\sin x}{x}+\frac{K}{\pi}\,\frac{t^2}{R^2}\left\{\frac{\sin x}{\gamma x}-\left(1+\frac{1}{\gamma}\cos x\right)\right\} + \text{terms of higher order in } \frac{1}{R}, \tag{18}$$

where x^2 is now $-k^2t^2\gamma$ and $\gamma=(d^2/k^2)-1$. To determine the limit of this expression it is necessary to normalize it, for the field strength should not be dependent upon R. Normalization may be accomplished by requiring the amplitude of the factor (18) at $r=R+a/2$ to be 1 for each value of R. Then $1/K=N_\nu(kR)J_\nu(kR+ka/2)-J_\nu(kR)N_\nu(kR+ka/2)$. Again by formula (2) of this paper, wherein we now let a be $a/2$, and by simple reductions and the taking of a limit as R becomes infinite, we find that expression (18) has the limiting form $L\sin(k^2-m^2\pi^2/b^2)^{1/2}(r-R)$, where L is independent of R and t. The k here is $k\big|_{R=\infty}$ which we have seen has the value $n^2\pi^2/a^2+m^2\pi^2/b^2$. This result shows that expression (18) permits us to exhibit the field in a coaxial sector as a function which approaches the field in the rectangular guide approached by the sector as R becomes infinite.

APPENDIX

This appendix will supply several theoretical details needed to complete the mathematics of the paper. In particular, it will justify the rearrangement of terms made in §2. This rearrangement converted a power series in α, with coefficients dependent upon ν and ρ, into a power series in ρ with coefficients dependent on k and d and subject to the condition that ν is a constant over ρ. As the body of the paper shows, when ν and ρ are replaced by their values in terms of R, the combination $\nu\rho$ is independent of R, the latter being the essential variable for the purpose of this paper.

Actual calculation of the quantities $\Delta_2, \Delta_3, \Delta_4, \cdots$, by operating with the basic recursion formula (12) of the paper, indicates that Δ_n for even n can be expressed as a polynomial in even powers of ρ with coefficients which are polynomials in even power of ν. Likewise, for odd n, Δ_n can be expressed as a polynomial in odd powers of ρ with coefficients which are polynomials in even powers of ν. More formally stated, we have the following theorem.

THEOREM A.

$$\Delta_{2n-1}(\rho) = \sum_{i=0}^{n-1} P_i^{2n-1}\, \rho^{2n-2i-1} \quad \textit{wherein} \quad P_i^{2n-1} = \sum_{j=0}^{n-1-i} a_{i,j}^{2n-1}\, \nu^{2(n-1-i-j)}$$

and

$$\Delta_{2n}(\rho) = \sum_{i=0}^{n-1} P_i^{2n}\, \rho^{2n-2i} \qquad \textit{wherein} \qquad P_i^{2n} = \sum_{j=0}^{n-1-i} a_{i,j}^{2n}\, \nu^{2(n-1-i-j)}.$$

Moreover, the coefficients $a_{i,j}^m$ satisfy the relations

(a)
$$a_{i,j}^{2n+1} = -(4n-4i-1)a_{i,j-1}^{2n} + a_{i,j}^{2n-1} - (2n-2i-1)^2 a_{i,j-1}^{2n-1} - a_{i-1,j}^{2n-1},$$

(b)
$$a_{i,j}^{2n+2} = -(4n-4i+1)a_{i,j}^{2n+1} + a_{i,j}^{2n} - (2n-2i-2)^2 a_{i,j-1}^{2n} - a_{i-1,j}^{2n},$$

provided that $a_{i,j}^m = 0$ when $i+j > (m-1)/2$, or $i<0$, or $j<0$, or $m<1$.

The proof of this theorem is readily made by mathematical induction utilizing the recursion formula (12).

It is now necessary to investigate the pattern of signs of the coefficients of the polynomials P_i^m with a view to showing that we are dealing with an absolutely convergent series. We next prove the following theorem.

THEOREM B. (1) *The sign of* $a_{i,j}^{m}=(-1)^{m+i+1}$, $m=1, 2, 3, \cdots$; $0\leqq i\leqq(m-1)/2$.

(2) *The sign of* $a_{i,j-1}^{2n-1}+(2n-2i-2)a_{i,j-2}^{2n-2}=(-1)^{i}$, $i+j\leqq n-1$.

(3) *The sign of* $a_{i,j-1}^{2n}+(2n-2i-1)a_{i,j-1}^{2n-1}=(-1)^{i+1}$, $i+j\leqq n-1$.

PROOF. It is (1) only, of Theorem B, which will be needed ultimately; however, the nature of the proof forces us to be concerned with parts (2) and (3). The proof is by mathematical induction. By inspection of Δ_{2n-2}, Δ_{2n-1}, and Δ_{2n}, when $n=1$ we find all three statements to be true. We now assume these statements true for $2n$, $2n-1$, and $2n-2$ and proceed to the main step of the inductive argument.

It follows from relation (a) of Theorem A, with j replaced by $j-1$ and the terms regrouped, that for all i, j

$$\text{(a)}\qquad \begin{aligned} &a_{i,j-1}^{2n+1}+(2n-2i)a_{i,j-2}^{2n}\\ &\quad=-(2n-2i-1)\{a_{i,j-2}^{2n}+(2n-2i-1)a_{i,j-2}^{2n-1}\}\\ &\qquad+a_{i,j-1}^{2n-1}-a_{i-1,j-1}^{2n-1}. \end{aligned}$$

By the induction assumption (3) the sign of the first term on the right is $(-1)^{i+2}=(-1)^{i}$; the sign of the second term, by assumption (1), is $(-1)^{2n-1+i+1}=(-1)^{i}$; and the sign of the third term is $(-1)^{1+(2n-1)+(i-1)+1}=(-1)^{i}$. It follows that assumption (2) holds for $n+1$.

By the induction assumption (1) the sign of $a_{i,j-2}^{2n}$ is $(-1)^{i+1}$ and the sign of the entire left side of (a) is $(-1)^{i}$ by the previous paragraph. It follows therefore that the term $a_{i,j-1}^{2n+1}$ on the left side of (a) has the sign $(-1)^{i}$. This step completes the inductive proof of statement (1) for the case where $m=2n+1$.

To establish statement (3) of the theorem we use relation (b) of Theorem A. If in that relation we replace j by $j-1$ and rearrange terms slightly, we obtain

$$\begin{aligned} &a_{i,j-1}^{2n+2}+(2n-2i+1)a_{i,j-1}^{2n+1}\\ &\quad=-(2n-2i)a_{i,j-1}^{2n+1}-(2n-2i-2)a_{i,j-2}^{2n}+a_{i,j-1}^{2n}-a_{i-1,j-1}^{2n}\\ &\quad=-(2n-2i)\{a_{i,j-1}^{2n+1}+(2n-2i)a_{i,j-2}^{2n}\}\\ &\qquad+2(4n-4i-2)a_{i,j-2}^{2n}+a_{i,j-1}^{2n}-a_{i-1,j-1}^{2n}. \end{aligned}$$

Since we have already established relation (2) for $n+1$, it follows that the sign of the first term on the right is $(-1)^{i+1}$. This is also the sign of each of the remaining terms in view of the inductive assumption for relation (1). Hence the right side and therefore the left side has the sign $(-1)^{i+1}$, which establishes relation (3) for $n+1$. At the

same time, since $a_{i,j-1}^{2n+1}$ has the sign $(-1)^i$ by the inductive proof of relation (1) when m is odd, it must be that the sign of $a_{i,j-1}^{2n+2}$ is $(-1)^{i+1}$ and this is the proof of relation (1) for $m=2n+2$, which completes the entire proof of that relation.

In consequence of (1) of Theorem B, the coefficients $a_{i,j}^m$ in the polynomials P_i^m in Theorem A are all of one sign, depending on m and i, and we obtain the following corollary.

COROLLARY I.

$$(-1)^i P_i^{2n-1} = \sum^{n-i-1} \left| a_{i,j}^{2n-1} \right| \nu^{2(n-1-i-j)},$$

$$(-1)^{i+1} P_i^{2n} = \sum_{j=1}^{n-i-1} \left| a_{i,j}^{2n} \right| \nu^{2(n-1-i-j)}.$$

We now let $\rho=(-1)^{1/2}r$ and substitute this and the results of Corollary I into the expressions for Δ_{2n-1} and Δ_{2n} of Theorem A. Further we let $\alpha=(-1)^{1/2}A$. Then we have the following corollary.

COROLLARY II. *For* $\alpha=(-1)^{1/2}A$ *and* $\rho=(-1)^{1/2}r$, $-\alpha^m\Delta_m(\rho) = A^m Q_m(r, \nu)$, $m=1, 2, 3, \cdots$, *where the coefficients in the polynomial* Q_m *are all positive.*

We are now near the conclusion. We recall that the Bessel functions $J_\nu(z)$ and $N_\nu(z)$ are analytic functions of the argument z everywhere except at $z=\infty$, and for the second of these functions at $z=0$. It follows at once that $G(\alpha, \beta)=J_\nu(\beta)N_\nu(\alpha+\beta)-J_\nu(\alpha+\beta)N_\nu(\beta)$ is analytic in α and β excepting possibly at the loci $\alpha=0$, $\alpha=\infty$, $\beta=0$, $\beta=\infty$, $\alpha+\beta=0$.

We already know from formula (5) that $(2\pi)G(\alpha, \rho^{-1}) = \sum_{m=1}^{\infty} (\alpha^m/m!)\Delta_m(\nu, \rho)$. The series on the right converges for all α and $\rho=\beta^{-1}$ which are not on the exceptional loci. By Corollary II, $(-2/\pi)G((-1)^{1/2}A, 1/(-1)^{1/2}r) = \sum_{m=1}^{\infty} (A^m/m!)Q_m(\nu, r)$. This power series in A converges for every A, for every fixed r, with exceptions as above, provided, of course, that in the summation we regard the terms in each $Q_m(\nu, r)$ as locked in a parenthesis, the entire polynomial serving as coefficient for A^m. But now, since each coefficient of $Q_m(\nu, r)$ is positive, we may regard the series on the right as a simple sequence of terms (A and r being thought of as fixed, arbitrary, real positive numbers) and we may certainly rearrange the terms of this simple sequence in any order, and reintroduce parentheses in any way we wish. The resulting series will converge to the same value as the original series.[6] Moreover r and A may also assume imaginary values

[6] See, for example, T. J. Bromwich, *An introduction to the theory of infinite series*, London, Macmillan, 1908, articles 31 and 33.

and the rearrangement of terms remains justified, for the absolute values of the imaginary r and A are the same as for real r and A and the convergence of the absolute values suffices to justify rearrangement.

An entirely analogous argument disposes of the same problem for the derivation of formula (16).

It should now be noted that the rearranged series, that is, formula (2), is a power series in $1/R$ which, because it equals the original series for $G(\alpha, \beta)$, certainly converges for all values of R except, perhaps, the few exceptional ones $R=0$, $R=\infty$, and $R=-a$. As a power series in $1/R$ it must also converge then for $R=\infty$ and it represents the function $G(\alpha, \beta)$ at $R=\infty$ if we define the function to have value 0 at that point. Formula (15) is therefore a Taylor's expansion of $G(\alpha, \beta)$ around $R=\infty$ under the condition $\nu=m\pi R/b$.[7]

NEW YORK UNIVERSITY

[7] Acknowledgment is hereby made of assistance from Dr. Leo Zippin in carrying out several inductive proofs of the paper.

ON A SIMPLE PROPERTY OF THE DERIVATIVES OF LEGENDRE'S POLYNOMIALS

EMIL GROSSWALD

It is well known that if we define, as usual, Legendre's polynomials as

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n, \tag{1}$$

we have

$$P_n(1) = 1 \qquad \text{for every } n. \tag{2}$$

More generally, for every n and r, the following relation holds:

$$\begin{aligned} P_n^{(r)}(1) &= C_{n+r,n-r}\cdot 1\cdot 3\cdots(2r-1) \\ &= \frac{(n-r+1)(n-r+2)\cdots(n+r)}{2\cdot 4\cdots 2r}. \end{aligned} \tag{3}$$

As we could find nowhere a proof of (3), except in the particular case $r=1$,[1] we shall give here a simple one.

From the classic relation[2]

$$P_n'(x) = nP_{n-1}(x) + x\cdot P_{n-1}'(x) \tag{4}$$

we obtain, by successive derivations,

$$P_n^{(r)}(x) = (n+r-1)\cdot P_{n-1}^{(r-1)}(x) + x\cdot P_{n-1}^{(r)}(x). \tag{4'}$$

Letting $x=1$ in (4), we obtain, for $n=\nu$, $P_\nu'(1)=\nu P_{\nu-1}(1)+P_{\nu-1}'(1)$, so that, by (2),

$$P_\nu'(1) = \nu + P_{\nu-1}'(1);$$

letting successively $\nu=1, 2, \cdots, n$ and adding, we obtain

$$P_n'(1) = \sum_{\nu=1}^{n}\nu = n(n+1)/2 \tag{5}$$

Received by the editors June 23, 1949.

[1] See Catalan, *Mémoire sur les fonctions X_n de Legendre*, Mémoires Couronnés de l'Academie de Belgique vol. 31 (1881).

[2] See D. Jackson, *Fourier series and orthogonal polynomials*, Carus Monographs, p. 48.

which proves (3) for $r=1$. Similarly, for $r=2$, we obtain from (4'), with $x=1$ and $n=\nu$,

$$P_\nu''(1) = (\nu+1)\cdot P_{\nu-1}'(1) + P_{\nu-1}''(1),$$

and, by (5),

$$P_\nu''(1) = (\nu+1)\frac{(\nu-1)\cdot\nu}{2} + P_{\nu-1}''(1).$$

Letting here, successively, $\nu=2, 3, \cdots, n$ and adding, we obtain:

$$P_n''(1) = \sum_{\nu=2}^{n} (\nu-1)\cdot\nu\cdot(\nu+1)/2 \tag{6}$$

$$= (n-1)\cdot n\cdot(n+1)(n+2)/2\cdot 4,$$

proving (3) for $r=2$. Let us assume now that (3) holds for all derivatives up to the order $r-1$; then, by (4'), we obtain, for $x=1$ and subscript ν:

$$P_\nu^{(r)}(1) = (\nu+r-1)P_{\nu-1}^{(r-1)}(1) + P_{\nu-1}^{(r)}(1), \tag{7}$$

or, using (3), valid by assumption for derivatives of order not greater than $r-1$,

$$P_\nu^{(r)}(1) = (\nu+r-1)\frac{(\nu-r+1)(\nu-r+2)\cdots(\nu+r-2)}{2\cdot 4\cdots(2r-2)} \tag{7'}$$

$$+ P_{\nu-1}^{(r)}(1).$$

Writing this relation for $\nu=r, r+1, \cdots, n-1, n$ and taking into account that $P_{r-1}^{(r)}(1)=0$, we obtain, adding,

$$P_n^{(r)}(1) = \sum_{\nu=r}^{n} \frac{(\nu-r+1)(\nu-r+2)\cdots(\nu+r-1)}{2\cdot 4\cdots(2r-2)}$$

$$= \frac{(n-r+1)(n-r+2)\cdots(n+r)}{2\cdot 4\cdots(2r)},$$

finishing the proof of (3).

From the form of (3) it results immediately, that all values of $P_n^{(r)}(1)$ are integers. As expressions containing $P_n^{(r)}(1)$ appear in connection with problems of divisibility, or irreducibility of polynomials, it may be of some interest to remark that, as a consequence of (3), $P_n^{(r)}(1)$ are always divisible by all odd factors of $(2r)!$, as well as by the highest power of 2 which divides $C_{n+r,n-r} = C_{n+r,2r}$.

UNIVERSITY OF PENNSYLVANIA

APPLICATIONS OF A LEMMA OF FEJÉR TO TYPICALLY-REAL FUNCTIONS

M. S. ROBERTSON

1. Introduction. Let

$$w = f(z) = z + a_2 z^2 + \cdots + a_n z^n + \cdots \tag{1.1}$$

be regular for $|z| < R$ and $f(x)$ real for $-R < x < R$. $f(z)$ is said to be typically-real with respect to the circle $|z| = R$ if the imaginary part of $f(z)$ has the same sign as the imaginary part of z when z is not real and $|z| < R$ [1].[1] In particular, if $f(z)$ is schlicht or univalent for $|z| < R$ and the coefficients a_n, $n = 2, 3, \cdots$, are all real, then $f(z)$ is a typically-real function. If further, $w = f(z)$ maps each circle $|z| = r < R$ into a contour having the property that every line parallel to the imaginary axis cuts the contour in not more than two points, we say that $f(z)$ is convex in the direction of the imaginary axis relative to the circle $|z| = R$. The necessary and sufficient condition that $f(z)$ be convex in the direction of the imaginary axis when it is real on the real axis is that $z f'(z)$ be typically-real [5; 4]. In the discussion to follow we shall assume that $R = 1$.

We state the following lemma due to L. Fejér (Turán [2], Szász [3]):

LEMMA 1. *For a value of* r, $0 < r < 1$, *let* $\sum_{n=1}^{\infty} \lambda_n r^n$ *converge. In order that*

$$\sum_{n=1}^{\infty} \lambda_n r^n \sin nx \cdot \sin ny \geqq 0 \qquad \text{for } 0 < x < \pi;\ 0 < y < \pi, \tag{1.2}$$

it is necessary and sufficient that

$$\sum_{n=1}^{\infty} n\lambda_n r^n \cdot \sin n\psi \geqq 0 \qquad \text{for } 0 < \psi < \pi. \tag{1.3}$$

It is the purpose of this paper to give some applications of Lemma 1 to functions

$$F(z) = \sum_{n=1}^{\infty} a_n b_n z^n \tag{1.4}$$

where

Presented to the Society, February 25, 1950; received by the editors June 27, 1949.

[1] Number in brackets refer to the references cited at the end of the paper.

$$f(z) = \sum_{n=1}^{\infty} a_n z^n, \qquad g(z) = \sum_{n=1}^{\infty} b_n z^n, \qquad a_1 = b_1 = 1, \tag{1.5}$$

are both typically-real with respect to the unit circle $|z| = 1$. The same method of proof may be used to obtain many special results for the functions $F(z)$ of (1.4). We include here the proofs of several theorems which illustrate the procedure.

THEOREM 1. *If*

$$f(z) = \sum_{n=1}^{\infty} a_n z^n, \qquad g(z) = \sum_{n=1}^{\infty} b_n z^n, \qquad a_1 = b_1 = 1, \tag{1.6}$$

are regular and typically-real for $|z| < 1$, *then*

$$F(z) = \sum_{n=1}^{\infty} a_n b_n z^n, \qquad |z| < 1, \tag{1.7}$$

is typically-real for $|z| \leqq 2 - 3^{1/2}$. *When*

$$f(z) \equiv g(z) \equiv z(1 - z)^{-2}, \qquad F(z) \equiv (z + z^2) \cdot (1 - z)^{-3},$$

$F(z)$ *is not typically-real for* $|z| < \rho$ *for any* $\rho > 2 - 3^{1/2}$.

THEOREM 2. *If*

$$f(z) = \sum_{n=1}^{\infty} a_n z^n, \qquad g(z) = \sum_{n=1}^{\infty} b_n z^n, \qquad a_1 = b_1 = 1, \tag{1.8}$$

are regular and typically-real for $|z| < 1$, *and if*

$$F(z) = \sum_{n=1}^{\infty} a_n b_n z^n, \tag{1.9}$$

then

$$h(z) \equiv \int_0^z \frac{F(z)}{z}\, dz = \sum_{n=1}^{\infty} \frac{a_n b_n}{n} z^n \tag{1.10}$$

is regular and typically-real for $|z| < 1$.

THEOREM 3. *If*

$$f(z) = \sum_{n=1}^{\infty} c_n z^n, \qquad g(z) = \sum_{n=1}^{\infty} d_n z^n, \qquad c_1 = d_1 = 1, \tag{1.11}$$

are regular, univalent, and convex in the direction of the imaginary axis for $|z| < 1$, *and real on the real axis, then*

$$F(z) = \sum_{n=1}^{\infty} c_n d_n z^n \tag{1.12}$$

is also regular, univalent, and convex in the direction of the imginary axis for $|z| < 1$.

If in Theorem 3 we take $g(z)$ to be the nth Cesàro partial sum of order 3 of the geometric series

$$z(1-z)^{-1} = z + z^2 + \cdots + z^n + \cdots, \tag{1.13}$$

which E. Egerváry [6] has shown to be univalent and convex for $|z| < 1$, we obtain immediately a new proof of the following theorem of L. Fejér [5].

THEOREM 4. (Fejér) *If*

$$f(z) = \sum_{n=1}^{\infty} c_n z^n \tag{1.14}$$

be regular, univalent, and convex in the direction of the imaginary axis for $|z| < 1$ *and real on the real axis, then the Cesàro partial sums of order 3 of* (1.14) *are also univalent and convex in the direction of the imaginary axis for* $|z| < 1$.

THEOREM 5. *If*

$$f(z) = \sum_{n=1}^{\infty} a_n z^n, \qquad g(z) = \sum_{n=1}^{\infty} b_n z^n, \qquad a_1 = b_1 = 1, \tag{1.15}$$

are regular and typically-real for $|z| < 1$, *and if*

$$F(z) = \sum_{n=1}^{\infty} a_n b_n z^n, \qquad |z| < 1 \tag{1.16}$$

then $\Re F(re^{i\theta}) \geqq 0$ *for* $|\theta| \leqq \arccos \mu(r) < \pi/2$, $0 < r < 1$, *where* $4r\mu(r) \equiv (1+34r^2+r^4)^{1/2} - 1 - r^2$. *The inequalities are sharp.*

Theorem 5 includes a result of W. Rogosinski [1] who has shown that $F(r) \geqq 0$, $0 < r < 1$. It is not hard to show from Theorem 5 that the function $f(z)$ of this theorem satisfies the inequality

$$\int_0^{2\pi} |f(ze^{i\theta}) \cdot f(z)| \cos \arg \left\{ \frac{f(ze^{i\theta})}{f(z)} \right\} d\phi \geqq 0, \tag{1.17}$$

for $z = re^{i\phi}$, $r < 1$, $|\theta| \leqq \arccos \mu(r^2)$, with equality holding for $\theta = \arccos \mu(r^2)$, $f(z) \equiv z(1-z)^{-2}$. Thus a specially weighted average of the cosine of the angle subtended at the origin by the points

$f(z), f(ze^{i\theta})$ is non-negative, although the cosine itself may be sometimes negative. We omit the proof of (1.17).

2. **Proof of Theorem 1.** It has been shown elsewhere [4] that the coefficients of the powers series for $f(z)$ and $g(z)$ in (1.5) have the representation

$$(2.1) \qquad a_n = \frac{1}{\pi}\int_0^\pi \frac{\sin nx}{\sin x}\,d\alpha(x), \qquad b_n = \frac{1}{\pi}\int_0^\pi \frac{\sin ny}{\sin y}\,d\beta(y)$$

where $\alpha(x)$, $\beta(y)$ are non-decreasing functions in the interval $(0, \pi)$ normalized so that

$$(2.2) \qquad \int_0^\pi d\alpha(x) = \int_0^\pi d\beta(x) = \pi.$$

It follows that the inequality

$$(2.3) \qquad IF(re^{i\theta}) = \sum_{n=1}^{\infty} a_n b_n r^n \sin n\theta \geqq 0$$

for $0<\theta<\pi$ for a suitable range of r can be deduced from the inequality

$$(2.4) \qquad \sum_{n=1}^{\infty} \frac{\sin nx}{\sin x}\cdot\frac{\sin ny}{\sin y}\cdot r^n \sin n\theta \geqq 0, \qquad 0 < \theta < \pi,$$

by term by term integrations with the aid of (2.1). However, if we apply Lemma 1 twice, we need only find the range of r for which $(z=re^{i\theta})$

$$(2.5) \qquad I\frac{(z+z^2)}{(1-z)^3} \equiv \sum_{n=1}^{\infty} n^2 r^n \sin n\theta \geqq 0, \qquad 0 < \theta < \pi.$$

$$\begin{aligned} I\{(z+z^2)\cdot(1-\bar{z})^3\} &= r\sin\theta\{1 - 6r^2 + r^4 - 2(r+r^3)\cos\theta\} \\ &\geqq r(1-r)^2\cdot(1+4r+r^2)\sin\theta \\ &\geqq 0 \quad \text{for } 0 \leqq \theta \leqq \pi,\ 0 \leqq r \leqq 2 - 3^{1/2}. \end{aligned} \tag{2.6}$$

The example $f(z)\equiv g(z)\equiv z(1-z)^{-2}$, $F(z)\equiv(z+z^2)(1-z)^{-3}$, shows that the range of r obtained above cannot be improved upon. This completes the proof of Theorem 1.

3. **Proofs of Theorems 2 and 3.** Using the notation of Theorem 2, we have

$$(3.1) \qquad h(z) = \sum_{n=1}^{\infty} \frac{a_n b_n}{n} z^n, \qquad |z| < 1,$$

$$(3.2)\qquad I\{h(z)\} = \sum_{n=1}^{\infty} \frac{a_n b_n}{n} r^n \sin n\theta, \qquad z = re^{i\theta}.$$

Thus $I\{h(z)\} \geqq 0$ for $0 \leqq \theta \leqq \pi$, $0 < r < 1$, follows from

$$(3.3)\qquad \sum_{n=1}^{\infty} \frac{1}{n} \cdot \frac{\sin nx}{\sin x} \cdot \frac{\sin ny}{\sin y} \cdot r^n \sin n\theta \geqq 0, \qquad 0 \leqq \theta \leqq \pi,$$

by term by term integrations and the aid of (2.1) and (2.2). But by Lemma 1, (3.3) may be replaced by

$$(3.4)\qquad \sum_{n=1}^{\infty} \frac{\sin n\psi}{\sin \psi} \cdot r^n \sin n\theta \geqq 0 \qquad \text{for } 0 < \psi < \pi,\ 0 \leqq \theta \leqq \pi.$$

$$(3.5)\qquad \sum_{n=1}^{\infty} \frac{\sin n\psi}{\sin \psi} r^n \sin n\theta = I\{z(1 - 2z\cos\psi + z^2)^{-1}\}.$$

Moreover, $z(1-2z\cos\psi+z^2)^{-1}$ is typically-real for $|z| < 1$, $0 \leqq \psi \leqq \pi$. Thus $h(z)$ is also typically-real for $|z| < 1$. This completes the proof of Theorem 2.

Theorem 3 follows from Theorem 2. For if

$$(3.6)\qquad f(z) = \sum_{n=1}^{\infty} c_n z^n, \qquad g(z) = \sum_{n=1}^{\infty} d_n z^n, \qquad c_1 = d_1 = 1,$$

be regular, univalent, and convex in the direction of the imaginary axis, and real on the real axis, then $zf'(z)$ and $zg'(z)$ are typically-real for $|z| < 1$. It follows from Theorem 2 that

$$(3.7)\qquad zF'(z) \equiv \sum_{n=1}^{\infty} \frac{(nc_n)(nd_n)}{n} z^n, \qquad |z| < 1,$$

is typically-real for $|z| < 1$. Thus

$$(3.8)\qquad F(z) \equiv \sum_{n=1}^{\infty} c_n d_n z^n, \qquad |z| < 1,$$

is regular, univalent, and convex in the direction of the imaginary axis.

4. **Proof of Theorem 5.** With the notation of Theorem 5 let

$$(4.1)\qquad F(z) \equiv \sum_{n=1}^{\infty} a_n b_n z^n, \qquad a_1 = b_1 = 1,\ |z| < 1.$$

With the aid of (2.1) and (2.2) it follows that the inequality

$$(4.2)\qquad \Re F(re^{i\theta}) \geqq 0, \qquad |\theta| \leqq \arccos \mu(r), \qquad 0 < r < 1,$$

can be deduced from the inequality

$$(4.3)\qquad \sum_{n=1}^{\infty} \frac{\sin nx}{\sin x}\cdot\frac{\sin ny}{\sin y}\cdot r^n \cos n\theta \geqq 0, \quad |\theta| \leqq \arccos \mu(r),$$

by the method used in the preceding paragraphs. By Lemma 1, we need only find when

$$(4.4)\qquad \sum_{n=1}^{\infty} n\frac{\sin n\psi}{\sin \psi}\cdot r^n \cos n\theta \geqq 0, \qquad 0 \leqq \psi \leqq \pi,$$

or, letting $\alpha=\pi-\theta$, $\phi=\pi-\psi$, when

$$(4.5)\qquad \sum_{n=1}^{\infty} n\frac{\sin n\phi}{\sin \phi}\, r^n \cos n\alpha \leqq 0, \qquad 0 \leqq \phi \leqq \pi,$$

or when

$$(4.6)\qquad R\{(z-z^3)\cdot(1-2z\cos\phi+z^2)^{-2}\} \leqq 0, \qquad z = re^{i\alpha}.$$

Since

$$(4.7)\qquad \begin{aligned} &R\{(\bar{z}-\bar{z}^3)\cdot(1-2z\cos\phi+z^2)^2\} \\ &\quad = r(r^2-1)[4r^2\cos^3\alpha - \{1+(6+4\cos^2\phi)r^2+r^4\}\cos\alpha \\ &\qquad + (4r+4r^3)\cos\phi], \end{aligned}$$

it will be sufficient to determine the values of α so that when $0<r<1$, $0\leqq\phi\leqq\pi$, we have

$$(4.8)\qquad \begin{aligned} &4r^2\cos^3\alpha - \{1+(6+4\cos^2\phi)r^2+r^4\}\cos\alpha \\ &\qquad + (4r+4r^3)\cos\phi \geqq 0. \end{aligned}$$

(4.8) may also be written in the form

$$(4.9)\qquad \begin{aligned} &(2r\cos\alpha - 1 - r^2)\{2r\cos^2\alpha + (1+r^2)\cos\alpha - 4r\} \\ &\qquad + 4r^2(1-\cos^2\phi)\cos\alpha - (1-\cos\phi)(4r+4r^3) \geqq 0; \end{aligned}$$

with regard to the last two terms of (4.9) we note that

$$(4.10)\qquad 4r^2(1-\cos^2\phi)\cos\alpha - (1-\cos\phi)(4r+4r^3) \geqq -8r(1+r^2)$$

with equality occurring in (4.10) for $\phi=\pi$ and for all α. Thus the left-hand side of inequality (4.8) is not less than

$$(4.11)\qquad 4r^2\cos^3\alpha - (1+10r^2+r^4)\cos\alpha - 4r - 4r^3$$

and the left-hand side of (4.8) equals the expression in (4.11) when $\phi=\pi$. Factoring (4.11) we have to determine when

$$(4.12)\qquad (1+2r\cos\alpha+r^2)\{2r\cos^2\alpha-(1+r^2)\cos\alpha-4r\}\geqq 0.$$

(4.12) holds when, and only when,

$$(4.13)\qquad \begin{aligned}-4r\leqq 4r\cos\alpha\leqq &-\{(1+34r^2+r^4)^{1/2}-1-r^2\}\\ &=-\mu(r)\cdot 4r.\end{aligned}$$

Thus $|\theta|=|\pi-\alpha|\leqq \arccos\mu(r)<\pi/2$.

The estimate for the bound on θ is sharp, as is seen by taking $\phi=\pi$, $f(z)\equiv g(z)\equiv z(1-z)^{-2}$.

References

1. W. Rogosinski, *Über positive harmonische Entwicklungen und typischreelle Potenzreihen*, Math. Zeit. vol. 35 (1932) pp. 93–121.

2. P. Turán, *Über die monotone Konvergenz der Cesàro-mittel bei Fourier- und Potenzreihen*, Proc. Cambridge Philos. Soc. vol. 34 (1938) pp. 134–143.

3. O. Szász, *On the partial sums of harmonic developments and of power series*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 12–21.

4. M. S. Robertson, *On the coefficients of a typically-real function*, Bull. Amer. Math. Soc. vol. 41 (1935) pp. 565–572.

5. L. Fejér, *Neue Eigenschaften der Mittelwerte bie den Fourierreihen*, J. London Math. Soc. vol. 8 (1933) pp. 53–62.

6. E. Egerváry, *Abbildungseigenschaften der arithmetischen Mittel der geometrischen Reihen*, Math. Zeit. vol. 42 (1937) pp. 221–230.

Rutgers University

ON CONFORMAL MAPPING OF NEARLY CIRCULAR REGIONS

S. E. WARSCHAWSKI

1. **Introduction.** The following problem has repeatedly been considered in the literature on conformal mapping. Given a simply connected region R which contains the origin and whose boundary lies between two concentric circles about the origin of radii 1 and $1+\epsilon$ ($\epsilon>0$) respectively. Suppose that the function $w=f(z)$ maps the circle $|z|<1$ conformally onto R in such a manner that $f(0)=0$ and $f'(0)>0$. It is desired to find: (a) an upper bound for the function $|f(z)-z|$ in terms of ϵ and r which holds for all z in a fixed circle $|z|\leqq r<1$, and (b) a bound which depends on ϵ only and is valid (possibly under additional assumption on the boundary of R) for all z in the complete (open or closed) unit circle. Estimates of the first kind were given by L. Bieberbach [1],[1] A. Ostrowski [10], C. Carathéodory [4 and 5], M. Lavrientieff and D. Kwasselava [7], and M. Müller [9]. Contributions to the second problem were made by L. Bieberbach [3], A. R. Marchenko [8], and J. Ferrand [6].

In the present paper we consider, more generally, *the integral means of order* $p>0$ of the function $|f(z)-z|$ taken along the circles $|z|=r$, namely

$$\mathfrak{M}_p\{f(z)-z\}=\left\{\frac{1}{2\pi}\int_0^{2\pi}|f(re^{i\theta})-re^{i\theta}|^p d\theta\right\}^{1/p}$$

and those of the first and second derivatives of this function. Bounds in terms of ϵ are obtained for these means which are valid uniformly for $r<1$. As an application of our estimate on the mean of the first derivative, we give a short proof of an inequality concerning $|f(z)-z|$ in $|z|\leqq 1$ which is due to A. R. Marchenko. Furthermore, a bound for the modulus of the derivative $f'(z)-1$ for z in $|z|\leqq 1$ is derived. In all cases the estimates obtained are of the form $O(\epsilon)$ as $\epsilon\to 0$ and are the best possible as far as the order of magnitude in ϵ is concerned.[2]

The knowledge of bounds for $|f(z)-z|$ and for some of the expressions considered in this paper is helpful in the application of certain iteration methods for determining the mapping function of a region onto a circle. For these inequalities may sometimes be used to ap-

Presented to the Society, April 28, 1944; received by the editors March 17, 1949.

[1] The numbers in the brackets refer to the bibliography at the end of the paper.

[2] This is easily shown by simple examples such as $f(z)=((z-c\epsilon)/(1-c\epsilon z)+\epsilon)(1/(1-c\epsilon))$ where c is a sufficiently small positive constant.

praise the accuracy of an approximation to the mapping function; see, for example, L. Bieberbach [1 and 2, p. 10], and a more recent paper of the writer [14]. In fact, the inequalities (5.1) and (6.1) for $p=2$ of the present note were stated and used in that paper (§9d) for such a purpose. However, the proofs of these statements were omitted there and are included in the present discussion of such inequalities.

The proofs of the inequalities involving the means make use of the following well known theorem of M. Riesz [11] on conjugate harmonic functions:

Suppose that the function $f(z)=u(z)+iv(z)$ *is regular for* $|z|<1$ *and that* $u(0)=0$. *Then for every* $p>1$ *there exists a positive number* A_p *which depends on* p *only, such that for* $0\leqq r<1$

$$(1.1)\quad \mathfrak{M}_p\{u(re^{i\theta})\} \equiv \left\{\frac{1}{2\pi}\int_0^{2\pi} |u(re^{i\theta})|^p\, d\theta\right\}^{1/p} \leqq A_p \mathfrak{M}_p\{v(re^{i\theta})\}.$$

If the right-hand side of (1.1) *is bounded for all* r, $0\leqq r<1$, *then* $f(re^{i\theta})$ *has radial boundary values of class* L_p *almost everywhere on* $|z|=1$ *and* (1.1) *holds for* $r=1$. *For* $p\geqq 2$ *all* $A_p\leqq 2p$ *and* $A_2=1$.

2. **The mean of the function** $f(z)-z$. We begin with the following theorem.

THEOREM I. *Let* R *be a simply connected region in the* w*-plane which contains the origin and whose boundary is contained in the ring*

$$(2.1)\qquad 1<|w|<1+\epsilon \qquad (\epsilon>0).$$

Suppose that the function $w=f(z)$ *maps the circle* $|z|<1$ *conformally onto* R *and that* $f(0)=0$ *and* $f'(0)>0$. *Then for every*[3] $p>1$ *and for* $z=re^{i\theta}$, *uniformly for* $0\leqq r<1$,

$$(2.2)\qquad \mathfrak{M}_p\{f(z)-z\} \leqq \mathfrak{M}_p\left\{\frac{f(z)}{z}-1\right\} \leqq (1+A_p)e^{\epsilon}\cdot\epsilon$$

where A_p *is the constant of the Riesz theorem.*[4]

[3] It is clear that the inequality (2.2) if proved for $p>1$ remains valid, for a suitable constant A_p, for all p, $0<p\leqq 1$ (Hölder's inequality). An analogous remark applies to the other estimates of the means derived in this paper, and we state and prove therefore Theorems I, II, III, and VI for $p>1$ only.

[4] By use of Cauchy's integral formula one may easily obtain from Theorem I a bound for $|f(z)-z|$ for all $|z|\leqq r<1$. If $t=\rho e^{i\theta}$, $r<\rho<1$, $|f(z)-z| \leqq 1/2\pi\int_0^{2\pi}|(f(t)-t)/(t-z)|\,d\theta \leqq \mathfrak{M}_p\{f(t)-t\}\,\mathfrak{M}_q\{(t-z)^{-1}\}$, by Hölder's inequality, where $1/p+1/q=1$. Hence, as $\rho\to 1$, $|f(z)-z|\leqq(1+A_p)e^{\epsilon}\epsilon\cdot(1-r)^{-1/p}$. However, this estimate is not the best possible as far as the order of $1/(1-r)$ is concerned, the best given in the literature being $\log 1/(1-r)$ (see [5, p. 51] and [9]).

PROOF. The function $F(z) \equiv f(z)/z$ for $0<|z|<1$, $F(0)=f'(0)$ is regular for $|z|<1$. Since the boundary of R is contained in the ring (2.1), a simple application of the principle of the maximum and minimum modulus shows that

$$1 \leqq |F(z)| \leqq 1+\epsilon \qquad \text{for } |z|<1.$$

Hence

$$0 \leqq \log |F(z)| \leqq \epsilon \qquad \text{for } |z|<1. \tag{2.3}$$

The branch of $\log F(z)=\log |F(z)|+i \arg F(z)$ for which $\arg F(0)=\arg f'(0)=0$ is single-valued and regular for $|z|<1$. By the theorem of M. Riesz stated above we obtain therefore for $z=re^{i\theta}$, $0 \leqq r<1$, and all $p>1$:

$$\mathfrak{M}_p\{\arg F(z)\} \leqq A_p \mathfrak{M}_p\{\log |F(z)|\} \leqq A_p \cdot \epsilon.$$

Hence, by use of Minkowski's inequality,

$$\mathfrak{M}_p\{\log F(z)\} \leqq \mathfrak{M}_p\{\log |F(z)|\} + \mathfrak{M}_p\{\arg F(z)\} \leqq (1+A_p)\epsilon. \tag{2.4}$$

Since for any complex number[5] a

$$|e^a - 1| \leqq |a|\, e^{|\Re a|},$$

we have (for $a=\log F(z)$) by (2.3)

$$\left|\frac{f(z)}{z} - 1\right| \leqq \left|\log \frac{f(z)}{z}\right| e^{\epsilon}$$

and therefore, by (2.4),

$$\mathfrak{M}_p\left\{\frac{f(z)}{z} - 1\right\} \leqq (1+A_p)e^{\epsilon} \cdot \epsilon.$$

3. **The mean of the function $zf'(z)/f(z)-1$.** Our next aim is to obtain an analogous result for the derivative of the mapping function. This will be accomplished essentially by the following theorem.

THEOREM II. *Let C be a closed Jordan curve represented (in polar coordinates) by the equation $\rho=\rho(\phi)$, $0 \leqq \phi \leqq 2\pi$, where $\rho(\phi)$ is positive, continuous, and moreover satisfies the inequality*

$$|\rho(\phi+\tau) - \rho(\phi)| \leqq \rho(\phi)\epsilon\,|\tau| \qquad (\textit{for some } \epsilon,\ 0<\epsilon<1). \tag{3.1}$$

Suppose that $w=f(z)$ maps the circle $|z|<1$ conformally onto the in-

[5] For, $|e^a-1| = |\int_0^a e^z dz| \leqq |\int_0^a |e^z|\,|dz|\,| \leqq |a|e^{|\Re a|}$, the integration being taken along the straight line segment $0 \cdots a$.

terior of C such that $f(0)=0$ and $f'(0)>0$. Then for all $p>1$ for which $\epsilon A_p<1$ and uniformly for $0<r<1$,

$$\mathfrak{M}_p\left\{z\frac{f'(z)}{f(z)}-1\right\} \leqq \frac{1+A_p}{1-\epsilon A_p}\epsilon \qquad (z=re^{i\theta}). \tag{3.2}$$

PROOF. Just as above let $F(z)$ be defined as $f(z)/z$ for $z\neq 0$, and as $f'(0)$ for $z=0$, and let $\log F(z)$ denote the branch of the logarithm which is real when $z=0$. Then $F(z)$ and $\log F(z)$ are regular in $|z|<1$ and since C is a Jordan curve, they are continuous in the closed circle $|z|\leqq 1$. Thus in particular $Q(z)\equiv \arg F(z)$ is single-valued and continuous in $|z|\leqq 1$. We define a branch of $\arg f(e^{i\theta})$ by the relation

$$\arg f(e^{i\theta}) = Q(e^{i\theta}) + \theta.$$

If $g(\theta)$ is a function defined for all real θ and δ is a positive number, we shall find it convenient to use the notation

$$\Delta g(\theta) = g(\theta+\delta) - g(\theta).$$

Since by hypothesis (3.1) $\log \rho(\phi)$ possesses bounded difference quotients, $\rho'(\phi)/\rho(\phi)$ exists almost everywhere and

$$|\log \rho(\phi+\tau) - \log \rho(\phi)| = \left|\int_{\phi}^{\phi+\tau}\frac{\rho'(t)}{\rho(t)}\,dt\right| \leqq \epsilon|\tau|.$$

Hence—note that $|f(e^{i\theta})| = \rho(\arg f(e^{i\theta}))$—

$$|\Delta \log |f(e^{i\theta})|| \leqq \epsilon|\Delta \arg f(e^{i\theta})|.$$

Now

$$\Delta \arg f(e^{i\theta}) = \Delta \arg F(e^{i\theta}) + \delta,$$

and therefore, by Minkowski's inequality, for $p>1$,

$$\mathfrak{M}_p\{\Delta \log |f(e^{i\theta})|\} \leqq \epsilon[\delta + \mathfrak{M}_p\{\Delta \arg F(e^{i\theta})\}]. \tag{3.3}$$

By the theorem of M. Riesz, applied to $\log F(e^{i\theta})$, we obtain

$$\mathfrak{M}_p\{\Delta \arg F(e^{i\theta})\} \leqq A_p\mathfrak{M}_p\{\Delta \log |F(e^{i\theta})|\}. \tag{3.4}$$

Hence, from (3.3)

$$\begin{aligned}\mathfrak{M}_p\{\Delta \log |F(e^{i\theta})|\} &= \mathfrak{M}_p\{\Delta \log |f(e^{i\theta})|\} \\ &\leqq \epsilon[\delta + A_p\mathfrak{M}_p\{\Delta \log |F(e^{i\theta})|\}]\end{aligned}$$

or

$$(1-\epsilon A_p)\mathfrak{M}_p\{\Delta \log |F(e^{i\theta})|\} \leqq \epsilon\delta. \tag{3.5}$$

Since $(1-\epsilon A_p)>0$, we may divide (3.5) by $1-\epsilon A_p$. We also divide by δ and, combining the so obtained inequality with (3.4), we find

$$\mathfrak{M}_p\left\{\frac{\Delta \log F(z)}{\delta}\right\} \leqq \frac{1+A_p}{1-\epsilon A_p}\epsilon \qquad (z=e^{i\theta}). \tag{3.6}$$

Since $\log F(z)$ is regular for $|z|<1$ and continuous for $|z|\leqq 1$, the inequality holds also when $z=re^{i\theta}$ where $0<r\leqq 1$. Replace in (3.6) z by $re^{i\theta}$ and keep $r<1$ *fixed*. The right-hand side of (3.6) is independent of δ. We may let $\delta\to 0$, passing to the limit under the integral sign, and thus we find (3.2).

4. **Applications of Theorem II.** As an application of the preceding result we obtain the desired estimate for the mean of the derivative.

THEOREM III. *Let C be a closed Jordan curve represented by the equation $\rho=\rho(\phi)$, $0\leqq\phi\leqq 2\pi$, where $\rho(\phi)$ is positive, continuous, and satisfies the following additional conditions: For some ϵ, $0<\epsilon<1$,*

(i) $$1 < \rho(\phi) < 1+\epsilon,$$

(ii) $$|\rho(\phi+\tau)-\rho(\phi)| \leqq \rho(\phi)\epsilon|\tau|.$$

If $w=f(z)$ is defined as in Theorem II, *then uniformly for $0\leqq r<1$, $p>1$, $\epsilon A_p<1$*

$$\mathfrak{M}_p\{f'(z)-1\} \leqq 2\frac{1+A_p}{1-\epsilon A_p}e^{\epsilon}\cdot\epsilon \qquad (z=re^{i\theta}). \tag{4.1}$$

To prove this we write for $0<|z|<1$

$$\begin{aligned}|f'(z)-1| &\leqq \left|f'(z)-\frac{f(z)}{z}\right|+\left|\frac{f(z)}{z}-1\right| \\ &\leqq \left|\frac{f(z)}{z}\right|\left|\frac{zf'(z)}{f(z)}-1\right|+\left|\frac{f(z)}{z}-1\right|.\end{aligned}$$

Since hypothesis (i) implies that $|f(z)/z|\leqq 1+\epsilon$ (see §2) we obtain by use of Minkowski's inequality:

$$\mathfrak{M}_p\{f'(z)-1\} \leqq (1+\epsilon)\mathfrak{M}_p\left\{\frac{zf'(z)}{f(z)}-1\right\}+\mathfrak{M}_p\left\{\frac{f(z)}{z}-1\right\}.$$

Hence by Theorems II and I:

$$\mathfrak{M}_p\{f'(z)-1\} \leqq (1+\epsilon)\frac{1+A_p}{1-\epsilon A_p}\epsilon+2\epsilon e^{\epsilon} \leqq \frac{2(1+A_p)}{1-\epsilon A_p}e^{\epsilon}\cdot\epsilon.$$

Theorem III is a sharpening of the following result which was first stated by A. R. Marchenko [8] in 1935 and which follows easily from (4.1).[6]

THEOREM IV. *Under the hypotheses of Theorem* III, *there exists an (absolute) constant* K *such that for* $|z| \leqq 1$

$$|f(z) - z| \leqq K\epsilon. \tag{4.2}$$

To prove this one needs only to apply the Fejér-Riesz inequality, which states that for any function $g(z)$ regular for $|z| \leqq r$,

$$\int_{-r}^{r} |g(\rho e^{i\alpha})|\, d\rho \leqq \frac{r}{2}\int_0^{2\pi} |g(re^{i\theta})|\, d\theta \quad (0 \leqq \alpha \leqq 2\pi).$$

For, if $z = re^{i\alpha}$, $0 < r < 1$,

$$|f(z) - z|^2 \leqq \left\{\int_0^r |f'(\rho e^{i\alpha}) - 1|\, d\rho\right\}^2 \leqq \int_0^r |f'(\rho e^{i\alpha}) - 1|^2 d\rho$$

$$\leqq \frac{r}{2}\int_0^{2\pi} |f'(re^{i\theta}) - 1|^2 d\theta \leqq \pi \cdot \frac{1}{2\pi}\int_0^{2\pi} |f'(re^{i\theta}) - 1|^2 d\theta$$

or

$$|f(z) - z| \leqq \pi^{1/2}\mathfrak{M}_2\{f'(z) - 1\} \leqq \frac{4e^{\epsilon}\pi^{1/2}}{1 - \epsilon}\epsilon \tag{4.3}$$

by (4.1)—note that $A_2 = 1$. Now, if $\epsilon \leqq 1/3$, the factor of ϵ in the last term of (4.3) does not exceed $6e^{1/3}\pi^{1/2}$, so that K in (4.3) may be taken equal to this number. If, however, $1/3 < \epsilon < 1$, then we make use of the fact that $|f(z) - z| \leqq 3 < 9\epsilon$ and choose $K = 9$. This proves (4.2).

5. **An estimate for the derivative of the mapping function.** The hypotheses of Theorem III are not sufficient to insure the existence of an inequality analogous to (4.2) for the derivative, $f'(z) - 1$ in the full circle $|z| < 1$, as under those assumptions $f'(z)$ need not even be bounded in $|z| < 1$. However, if we add a further condition (essentially involving the curvature of C) then we can prove the following:

THEOREM V. *Let* C *be a closed Jordan curve which satisfies hypotheses*

[6] Marchenko stated his theorem without proof. Another proof of this inequality, essentially based on the use of the Poisson integral, is given by E. J. Specht in a forthcoming paper [12]. In addition, bounds in terms of ϵ for the derivatives $|f'(z) - 1|$, $|f''(z)|$, $\cdots$, $|f^{(n)}(z)|$ which are valid in $|z| \leqq 1$ are derived in this paper. Specht's estimate for the first derivative is based on somewhat less restrictive assumptions than those of our Theorem V below, and yields a bound of the same order of magnitude in ϵ, if applied under the hypotheses of our Theorem V.

(i) *and* (ii) *of Theorem* III. *Furthermore, suppose that* $\rho'(\phi)$ *exists everywhere and that*[7]

$$\text{(iii)} \qquad \left|\frac{\rho'}{\rho}(\phi+\tau)-\frac{\rho'}{\rho}(\phi)\right|<\epsilon|\tau|,$$

ϵ *being the same as in* (i) *and* (ii). *Suppose that* $f(z)$ *is defined as in Theorem* III *and that* $\phi(\theta)=\arg f(e^{i\theta})$. *Then* $f'(z)$ *exists and is continuous in* $|z|\leqq 1$ *and*

$$\text{(5.1)} \qquad \frac{1}{A(1+\epsilon^2)^{1/2}}\leqq \phi'(\theta)\leqq A, \qquad 0\leqq\theta\leqq 2\pi,$$

where $A=4\cdot e^{\epsilon^2}$ *and*

$$\text{(5.2)} \qquad |f'(z)-1|\leqq 2B(1+\epsilon)\cdot\epsilon+\epsilon, \qquad |z|\leqq 1,$$

where $B=A+(A-1)/\epsilon$.[8]

Proof of (5.1). Under the present hypotheses the mapping function $f(z)$ possesses a continuous nonvanishing derivative[9] $f'(z)$ for $|z|\leqq 1$. If $\log(f(z)/z)$ again denotes the branch of the logarithm which reduces to the real $\log f'(0)$ when $z=0$, then we have, for $z=e^{i\theta}$,

$$\text{(5.3)} \qquad \log\frac{f(e^{i\theta})}{e^{i\theta}}=\log\rho[\phi(\theta)]+i\{\phi(\theta)-\theta\}.$$

Differentiation with respect to θ leads to the relation

$$\text{(5.4)} \qquad \left\{\frac{f'(z)}{f(z)}-\frac{1}{z}\right\}iz=\frac{\rho'}{\rho}[\phi(\theta)]\phi'(\theta)+i\{\phi'(\theta)-1\}$$

or

$$\text{(5.5)} \qquad \left\{\frac{zf'(z)}{f(z)}\right\}_{z=e^{i\theta}}=\phi'(\theta)-i\frac{\rho'}{\rho}[\phi(\theta)]\phi'(\theta).$$

---

[7] The notation $(\rho'/\rho)(\phi)$ means $\rho'(\phi)/\rho(\phi)$.

[8] In [14, § 9d], the hypothesis on $\rho(\phi)$ which corresponds to condition (i) of Theorem V of the present paper is $(1+\epsilon)^{-1}a\leqq\rho(\phi)\leqq a(1+\epsilon)$, where a is a positive constant. The constant $A(=e^{\epsilon^2}4^{\epsilon})$ is given there erroneously for the case proved here under the assumption (i), and it should be replaced by $A=e^{2\epsilon^2}4^{\epsilon}$. This is seen by an obvious modification of our proof of (5.1); only the relations (5.10) and (5.11) and the inequalities which follow from them must be changed.

[9] See, for example, [13, p. 327].

Since $f'(e^{i\theta}) \neq 0$, it follows that $\phi'(\theta) \neq 0$. The function

$$\log \frac{zf'(z)}{f(z)} = \log \left| \frac{zf'(z)}{f(z)} \right| + i \arg \frac{zf'(z)}{f(z)} \tag{5.6}$$

which is defined as being 0 when $z=0$, is single-valued and analytic for $|z|<1$, and continuous for $|z| \leqq 1$. If we set

$$\omega(\phi) = -\arctan \frac{\rho'}{\rho}[\phi],$$

where the principal value of the arctangent is chosen, then, for $z=e^{i\theta}$,

$$\arg \frac{zf'(z)}{f(z)} = \omega[\phi(\theta)], \quad \log \left| \frac{zf'(z)}{f(z)} \right| = \log \{\phi'(\theta) \sec \omega[\phi(\theta)]\}. \tag{5.7}$$

By a well known formula for the conjugate function we have therefore

$$\log \{\phi'(\theta) \sec \omega[\phi(\theta)]\} = \frac{1}{2\pi} \int_0^\pi \{\omega[\phi(\theta+\tau)] - \omega[\phi(\theta-\tau)]\} \cot \frac{\tau}{2} d\tau.$$

Because of hypothesis (iii) it follows easily that

$$|\omega[\phi(\theta+\tau)] - \omega[\phi(\theta-\tau)]| \leqq \epsilon\{\phi(\theta+\tau) - \phi(\theta-\tau)\}.$$

Hence

$$|\log \{\phi'(\theta) \sec \omega[\phi(\theta)]\}| \leqq \frac{\epsilon}{2\pi} \int_0^\pi \{\phi(\theta+\tau) - \phi(\theta-\tau)\} \cot \frac{\tau}{2} d\tau,$$

or

$$|\log \{\phi'(\theta) \sec \omega[\phi(\theta)]\}| \leqq \frac{\epsilon}{2\pi} \int_0^\pi \{\phi(\theta+\tau) - (\theta+\tau) - [\phi(\theta-\tau) - (\theta-\tau)]\} \cot \frac{\tau}{2} d\tau + \frac{\epsilon}{\pi} \int_0^\pi \tau \cot \frac{\tau}{2} d\tau. \tag{5.8}$$

The second term on the right-hand side of (5.8) has the value $2\epsilon \log 2$. Furthermore, since $\log \rho[\phi(\theta)]$ and $\phi(\theta)-\theta$ are real and imaginary parts of $\log (f(z)/z)$ for $z=e^{i\theta}$, and the real part of this function is $\log f'(0)$ when $z=0$, it follows, again by the formula for the conjugate function, that the first term on the right-hand side of (5.8) is equal to

$\epsilon\{\log \rho[\phi(\theta)] - \log f'(0)\}$. Thus

$$|\log\{\phi'(\theta)\sec\omega[\phi(\theta)]\}| \leqq \epsilon \log \frac{\rho[\phi(\theta)]}{f'(0)} + 2\epsilon\log 2. \tag{5.9}$$

By hypothesis (i) we have for $|z|=1$ and therefore also for $0<|z|<1$

$$1 \leqq \left|\frac{f(z)}{z}\right| \leqq 1+\epsilon \tag{5.10}$$

and thus, as $z\to 0$,

$$1 \leqq f'(0) \leqq 1+\epsilon. \tag{5.11}$$

We have therefore

$$\log \frac{\rho[\phi(\theta)]}{f'(0)} \leqq \log(1+\epsilon) < \epsilon.$$

Thus, we finally find from (5.9)

$$-\log\sec\omega[\phi(\theta)] - \epsilon^2 - 2\epsilon\log 2 \leqq \log\phi'(\theta) \leqq \epsilon^2 + 2\epsilon\log 2. \tag{5.12}$$

By hypothesis (ii)

$$\begin{aligned} -\log\sec\omega[\phi(\theta)] &= \frac{1}{2}\log\frac{1}{1+((\rho'/\rho)[\phi(\theta)])^2} \\ &\geqq \frac{1}{2}\log\frac{1}{1+\epsilon^2}. \end{aligned} \tag{5.13}$$

The inequalities (5.12) and (5.13) establish (5.1).

PROOF OF (5.2). From (5.4), hypothesis (ii), and (5.1) we obtain

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leqq \left|\frac{\rho'}{\rho}[\phi(\theta)]\right| |\phi'(\theta)| + |\phi'(\theta)-1| \leqq \epsilon A + |\phi'(\theta)-1|.$$

A simple calculation shows that $1-1/A(1+\epsilon^2)^{1/2} \leqq A-1$, so that, by (5.1), $|\phi'(\theta)-1| \leqq A-1$ and therefore

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leqq \epsilon A + A - 1. \tag{5.14}$$

Next we write

$$|f'(z)-1| \leqq \left|f'(z) - \frac{f(z)}{z}\right| + \left|\frac{f(z)}{z} - f'(0)\right| + |f'(0)-1|$$

$$\leqq \left|\frac{f(z)}{z}\right|\left|\frac{zf'(z)}{f(z)} - 1\right| + \int_0^1 \left|\frac{d}{dt}\left(\frac{f(t)}{t}\right)\right| |dt| + |f'(0) - 1|.$$

Using (5.10), (5.14), and (5.11) we find

$$|f'(z) - 1| \leqq (1+\epsilon)(A\epsilon + A - 1) + \epsilon + \int_0^1 \left|\frac{d}{dt}\left(\frac{f(t)}{t}\right)\right| |dt|. \tag{5.15}$$

Now the function

$$\frac{d}{dt}\left(\frac{f(t)}{t}\right) = \frac{f(t)}{t}\cdot\frac{1}{t}\cdot\left(\frac{tf'(t)}{f(t)} - 1\right)$$

is regular for $|t| < 1$ (including $t = 0$, if properly defined) and continuous for $|t| \leqq 1$. But for $|t| = 1$ we have by (5.10) and (5.14)

$$\left|\frac{d}{dt}\left(\frac{f(t)}{t}\right)\right| \leqq (1+\epsilon)(A\epsilon + A - 1),$$

and by the principle of the maximum modulus this holds also for $|t| < 1$. Combining this result with (5.15) we obtain (5.2).

6. **The mean of the second derivative of the mapping function.** The assumptions made in Theorem V permit one to go beyond the statement proved on the first derivative of the mapping function and to obtain estimates for the mean of order p of the second derivative.

THEOREM VI. *Under the hypotheses of Theorem* V, $\phi'(\theta)$ *is an absolutely continuous function and for any* $p \geqq 2$

$$\mathfrak{M}_p\{\phi''(\theta)\} \leqq AM_p\epsilon \tag{6.1}$$

where $A = 4e^{\epsilon}\epsilon^2$ *and* $M_p = A_p A^{1-1/p} \min(1+\epsilon, 2^{1-1/p})$, *and for* $p > 1$

$$\mathfrak{M}_p\left(\frac{f''(z)}{f'(z)}\right) \leqq (A + B + AA_p)\epsilon, \qquad (z = re^{i\theta}, 0 \leqq r < 1) \tag{6.2}$$

where $B = A + (A-1)/\epsilon$.

PROOF OF (6.2). We consider as in equation (5.6) the branch of $\log(zf'(z)/f(z))$ which is defined as 0 when $z = 0$. Then $\arg(zf'(z)/f(z))$ is harmonic for $|z| < 1$ and continuous for $|z| \leqq 1$ and by (5.7)

$\arg (zf'(z)/f(z)) = \omega[\phi(\theta)]$ for $z = e^{i\theta}$. Because of hypothesis (iii) and (5.1), $\omega[\phi(\theta)]$ possesses bounded difference quotients for all θ, so that $\omega'[\phi(\theta)]\phi'(\theta)$ exists for almost all θ and is bounded. This implies that the function $\partial \arg (zf'(z)/f(z))/\partial\theta$ may be represented by the Poisson integral for $|z| = |re^{i\theta}| < 1$. For, if we write

$$K(r, \theta - t) = \frac{1 - r^2}{1 + r^2 - 2r\cos(\theta - t)},$$

we have

$$\begin{aligned}\frac{\partial}{\partial\theta}\left[\arg\frac{zf'(z)}{f(z)}\right] &= \frac{\partial}{\partial\theta}\frac{1}{2\pi}\int_0^{2\pi}\omega[\phi(t)]K(r, \theta - t)dt \\ &= -\frac{1}{2\pi}\int_0^{2\pi}\omega[\phi(t)]\frac{\partial}{\partial t}K(r, \theta - t)dt \\ &= \frac{1}{2\pi}\int_0^{2\pi}\omega'[\phi(t)]\phi'(t)K(r, \theta - t)dt,\end{aligned}$$

and from this representation it follows that, for $z = re^{i\theta}$,

$$\mathfrak{M}_p\left\{\frac{\partial}{\partial\theta}\left[\arg\frac{zf'(z)}{f(z)}\right]\right\} \leqq \mathfrak{M}_p\{\omega'[\phi(\theta)]\phi'(\theta)\} \leqq \epsilon A^{(p-1)/p}, \tag{6.3}$$

by hypothesis (iii) and (5.1).[10]

We apply now the theorem of M. Riesz to $\partial \log (zf'(z)/f(z))/\partial\theta$ and find from (6.3) that, for $z = re^{i\theta}$, $0 < r < 1$,

$$\mathfrak{M}_p\left\{\frac{\partial}{\partial\theta}\log\left|\frac{zf'(z)}{f(z)}\right|\right\} \leqq A_p\mathfrak{M}_p\{\omega'[\phi(\theta)]\phi'(\theta)\} \leqq A_pA^{1-1/p}\cdot\epsilon, \tag{6.4}$$

and therefore

$$\mathfrak{M}_p\left\{\frac{\partial}{\partial\theta}\log\frac{zf'(z)}{f(z)}\right\} \leqq \epsilon A^{1-1/p}(1 + A_p). \tag{6.5}$$

Since

$$\frac{\partial}{\partial\theta}\log\frac{zf'(z)}{f(z)} = iz\frac{f''(z)}{f'(z)} - i\left(z\frac{f'(z)}{f(z)} - 1\right),$$

it follows from (6.5) and (5.14) that

$$\mathfrak{M}_p\left\{\frac{zf''(z)}{f'(z)}\right\} \leqq \epsilon A^{1-1/p}(1 + A_p) + B\epsilon \leqq (A + B + AA_p)\epsilon.$$

This proves (6.2).

[10] Note that $\{(1/2\pi)\int_0^{2\pi}(\phi')^p d\theta\}^{1/p} \leqq \{(A^{p-1}/2\pi)\int_0^{2\pi}\phi' d\theta\}^{1/p} = A^{(p-1)/p}$.

PROOF OF (6.1). The inequality (6.4) implies that the function $\log\left|zf'(z)/f(z)\right|$ for $z=e^{i\theta}$ is an absolutely continuous function of θ. Hence, by the second of the relations (5.7), $\phi'(\theta)$ is an absolutely continuous function. Furthermore $\lim_{r\to 1}\partial\log\left|rf'(re^{i\theta})/f(re^{i\theta})\right|/\partial\theta$ exists for almost all θ and is equal to

$$\phi''(\theta)/\phi'(\theta)+\tan\omega[\phi(\theta)]\omega'[\phi(\theta)]\phi'(\theta).$$

We may pass to the limit as $r\to 1$ under the integral sign in (6.4), thus obtaining

$$\mathfrak{M}_p\left\{\frac{\phi''}{\phi'}+\tan\omega\cdot\omega'\cdot\phi'\right\}\leqq A_p\mathfrak{M}_p\{\omega'\cdot\phi'\}. \tag{6.6}$$

In order to estimate $\mathfrak{M}_p\{\phi''/\phi'\}$ we write

$$\phi''/\phi'=(\phi''/\phi'+\tan\omega\cdot\omega'\cdot\phi')-\tan\omega\cdot\omega'\cdot\phi'$$

and apply Minkowski's inequality. Using (6.6) we find

$$\mathfrak{M}_p\left\{\frac{\phi''}{\phi'}\right\}\leqq A_p\mathfrak{M}_p\{\omega'\phi'\}+\mathfrak{M}_p\{\tan\omega\cdot\omega'\cdot\phi'\}, \tag{6.7}$$

and by hypotheses (ii) and (iii) and (5.1)

$$\mathfrak{M}_p\left\{\frac{\phi''}{\phi'}\right\}\leqq A_pA^{1-1/p}\epsilon+A^{1-1/p}\epsilon^2=A^{1-1/p}(A_p+\epsilon)\epsilon. \tag{6.8}$$

We can estimate the right-hand side of (6.7) in a different way. First, we note that it does not exceed[11]

$$\left\{\frac{1}{2\pi}\int_0^{2\pi}2^{p-1}\left|\omega'\cdot\phi'\right|^p(A_p^p+\left|\tan\omega\right|^p)d\theta\right\}^{1/p}$$

$$\leqq A_p2^{1-1/p}\left\{\frac{1}{2\pi}\int_0^{2\pi}\left|\omega'\cdot\phi'\right|^p(1+\left|\tan\omega\right|^p)d\theta\right\}^{1/p}.$$

Next, using the fact that

$$\left|\omega'(\phi)\right|=\left|-\frac{\dfrac{d}{d\phi}\left(\dfrac{\rho'}{\rho}\right)}{1+\left(\dfrac{\rho'}{\rho}\right)^2}\right|\leqq\frac{\epsilon}{1+\left(\dfrac{\rho'}{\rho}\right)^2}$$

we obtain

$$\mathfrak{M}_p\left\{\frac{\phi''}{\phi'}\right\}\leqq\epsilon A_p2^{1-1/p}\left\{\frac{1}{2\pi}\int_0^{2\pi}(\phi')^p\,\frac{1+\left|\rho'/\rho\right|^p}{(1+(\rho'/\rho)^2)^p}\,d\theta\right\}^{1/p}.$$

[11] We are making use here of the fact that, for $a\geqq 0$, $b\geqq 0$, $p\geqq 1$, $a^{1/p}+b^{1/p}\leqq 2((a+b)/2)^{1/p}=(2^{p-1}(a+b))^{1/p}$.

Now since $p \geqq 2$

$$1 + \left|\frac{\rho'}{\rho}\right|^p = 1 + \left[\left(\frac{\rho'}{\rho}\right)^2\right]^{p/2} \leqq \left(1 + \left(\frac{\rho'}{\rho}\right)^2\right)^p$$

and therefore

$$\mathfrak{M}_p\left\{\frac{\phi''}{\phi'}\right\} \leqq 2^{1-1/p} A_p A^{1-1/p} \epsilon. \tag{6.9}$$

Thus we find from (6.8) and (6.9) that

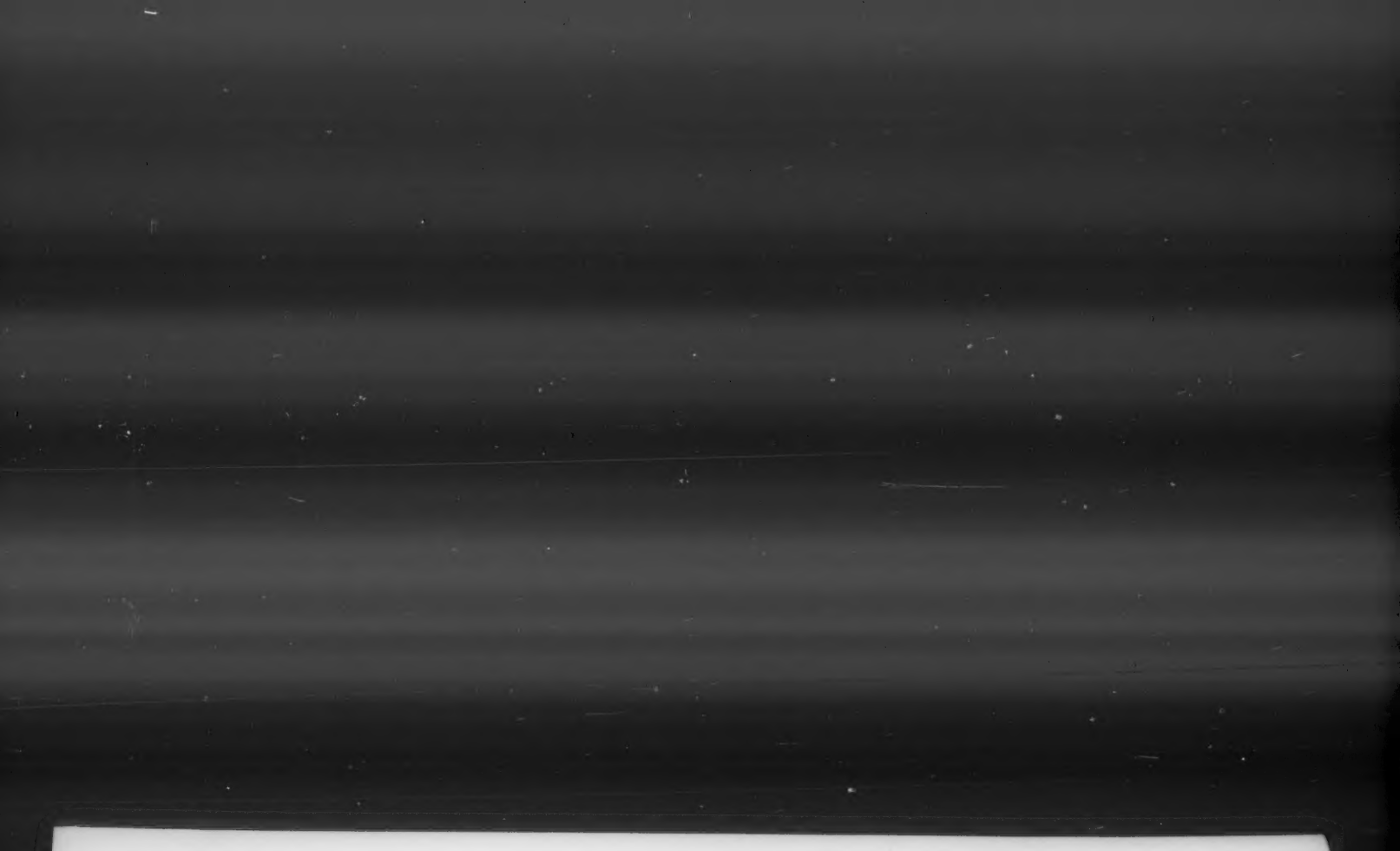
$$\mathfrak{M}_p\left\{\frac{\phi''}{\phi'}\right\} \leqq A_p A^{1-1/p} \epsilon \min\,(1+\epsilon,\, 2^{1-1/p}).$$

Finally, if we note again that $\phi'(\theta) \leqq A$, we obtain (6.1).

Bibliography

1. L. Bieberbach, *Zur Theorie und Praxis der konformen Abbildung*, Rend. Circ. Mat. Palermo vol. 33 (1914) pp. 93–112.
2. ———, *Lehrbuch der Funktionentheorie*, vol. 2, 2d ed., Teubner, Leipzig, 1931.
3. ———, *Über die Kreisabbildung von nahezu kreisförmigen Bereichen*, Preuss. Akad. Wiss. Sitzungsber. (1924) pp. 181–188.
4. C. Carathéodory, *Elementarer Beweis für den Fundamentalsatz der konformen Abbildung*, Schwarz-Festschrift, Springer, Berlin, 1914, pp. 19–41 (part 7).
5. ———, *Conformal representation*, Cambridge University Tracts, no. 28, 1932.
6. J. Ferrand, *Sur la déformation analytique d'un domain*, C. R. Acad. Sci. Paris vol. 221 (1945) pp. 132–134.
7. M. Lavrientieff and D. Kwasselava, *Über einen Ostrowskischen Satz*, Georg. Abt. Akad. d. Wiss., USSR vol. 1 (1940) pp. 171–174.
8. A. R. Marchenko, *Sur la représentation conforme*, C. R. (Doklady) Acad. Sci. USSR vol. 1 (1935) pp. 289–290.
9. M. Müller, *Zur konformen Abbildung angenähert kreisförmiger Gebiete*, Math. Zeit. vol. 43 (1938) pp. 628–636.
10. A. Ostrowski, *Über konforme Abbildungen annähernd kreisförmiger Gebiete*, Jber. Deutschen Math. Verein. vol. 39 (1930) pp. 78–81.
11. M. Riesz, *Sur les fonctions conjuguées*, Math. Zeit. vol. 27 (1928) pp. 218–244.
12. E. J. Specht, *Estimates on the mapping function and its derivatives in conformal mapping of nearly circular regions*, to be published.
13. S. E. Warschawski, *On the higher derivatives at the boundary in conformal mapping*, Trans. Amer. Math. Soc. vol. 38 (1935) pp. 310–340.
14. ———, *On Theodorsen's method of conformal mapping of nearly circular regions*, Quarterly of Applied Mathematics vol. 3 (1945) pp. 12–28.

University of Minnesota



Books Published by the

American Mathematical Society

Colloquium Publications

35. A. C. Schaeffer and D. C. Spencer, **Coefficient regions for schlicht functions**, to appear soon. $6.00
34. J. L. Walsh, **The location of critical points of analytic and harmonic functions**, 1950, viii, 384 pp. 6.00
33. J. F. Ritt, **Differential algebra**, 1950, viii, 184 pp. 4.40
32. R. L. Wilder, **Topology of manifolds**, 1949, x, 402 pp. 7.00
31. E. Hille, **Functional analysis and semi-groups**, 1948, xii, 528 pp. 7.50
30. T. Radó, **Length and area**, 1948, vi, 572 pp. 6.75
29. A. Weil, **Foundations of algebraic geometry**, 1946, xx, 288 pp. 5.50
28. G. T. Whyburn, **Analytic topology**, 1942; reprinted, 1948, x, 280 pp. 4.75
27. S. Lefshetz, **Algebraic topology**, 1942; reprinted, 1948, vi, 393 pp. 6.00
26. N. Levinson, **Gap and density theorems**, 1940, viii, 246 pp. 4.00
25. G. Birkhoff, **Lattice theory**, 1940; enlarged and completely revised edition, 1949, xiv, 283 pp. 6.00
24. A. A. Albert, **Structure of algebras**, 1939, xii, 210 pp. 4.00
23. G. Szegö, **Orthogonal polynomials**, 1939; reprinted, 1949, x, 403 pp. 6.00
22. C. N. Moore, **Summable series and convergence factors**, 1938, vi, 105 pp. 2.00
21. J. M. Thomas, **Differential systems**, 1937, x, 118 pp. 2.00
20. J. L. Walsh, **Interpolation and approximation by rational functions in the complex domain**, 1935, x, 382 pp. 6.00
19. R. E. A. C. Paley and N. Wiener, **Fourier transforms in the complex domain**, 1934; reprinted, 1949, viii, 184 pp. +portrait plate 3.75
18. M. Morse, **The calculus of variations in the large**, 1934; reprinted, 1948, x, 368 pp. 5.75
17. J. H. M. Wedderburn, **Lectures on matrices**, 1934; reprinted, 1946; reprinted, 1949, x, 205 pp. 3.35
16. G. A. Bliss, **Algebraic functions**, 1933; reprinted, 1948, x, 220 pp. 3.75
14. J. F. Ritt, **Differential equations from the algebraic standpoint**, 1932; reprinted, 1948, x, 172 pp. 3.00
11. D. Jackson, **The theory of approximation**, 1930; reprinted, 1946, viii, 178 pp. 3.35
10. A. B. Coble, **Algebraic geometry and theta functions**, 1929; reprinted, 1948, viii, 282 pp. 4.00
9. G. D. Birkhoff, **Dynamical systems**, 1927; reprinted, 1948, viii, 295 pp. 4.25
8. L. P. Eisenhart, **Non-Riemannian geometry**, 1927; reprinted, 1949, viii, 184 pp. 2.70
5. O. Veblen, **Analysis situs**, 1922; second ed., 1931; reprinted, 1946, x, 194 pp. 3.35
3. E. Kasner, **Differential-geometric aspects of dynamics**, 1913; reprinted, 1934; reprinted, 1948, iv, 117 pp. 2.50
1. H. S. White, **Linear systems of curves on algebraic surfaces**;
   F. S. Woods, **Forms of non-Euclidean space**;
   E. B. Van Vleck, **Selected topics in the theory of divergent series and of continued fractions**; 1905, xii, 187 pp. 3.00

Mathematical Surveys

4. O. F. G. Schilling, **The theory of valuations**, viii, 253 pp. $6.00
3. M. Marden, **The geometry of the zeros of a polynomial in a complex variable**, 1949, x, 181 pp. 5.00
2. N. Jacobson, **The theory of rings**, 1943; reprinted, 1949, vi, 151 pp. 3.35

Semicentennial Publications

1. R. C. Archibald, **A semicentennial history of the American Mathematical Society 1888–1938, with biographies and bibliographies of the past presidents**, 1938, xii, 261 pp. +30 plates. $3.00
2. E. T. Bell, J. F. Ritt, N. Wiener, E. J. McShane, T. Y. Thomas, R. L. Wilder, G. C. Evans, J. L. Synge, and G. D. Birkhoff, **Semicentennial addresses**, 1938, vi, 315 pp. 4.50

Proceedings of Symposia in Applied Mathematics

2. **Electromagnetic theory**, 1950, iv, 91 pp. $5.00
1. **Non-linear problems in mechanics of continua**, 1949, iv, 219 pp. 5.25

The price to members of the American Mathematical Society is 25% less than list
Volumes which are not listed above are out of print.

CONTENTS

Vol. 1, No. 4 AUGUST, 1950 Whole No. 4

Page

On Chern's invariant for Riemannian 4-manifolds. By HANS SAMELSON 415
Finitely-valued cohomology groups. By J. W. KEESEE 418
A note on a combinatorial problem. By H. J. RYSER 422
Linear representations of topological groups. By MORIKUNI GOTÔ 425
On a conjecture on simple groups. By I. N. HERSTEIN 438
On a semi-group of subsets of a linear space. By R. E. FULLERTON 440
Note on the arithmetic of bilinear transformations. By D. M. ADELMAN 443
Minimal identities for algebras. By A. S. AMITSUR and J. LEVITZKI 449
On matrices whose characteristic equations are identical. By W. V. PARKER 464
Simply connected homogeneous spaces. By DEANE MONTGOMERY 467
Roots and factors of ordinals. By P. W. CARRUTH 470
On the decomposition of orthogonalities into symmetries. By PETER SCHERK 481
On the polynomial solutions of the generalized Lamé differential equation. By MORRIS MARDEN 492
On the algebraic integrals of a system of differential equations of mechanics. By J. A. BREVES FILHO 498
Note on the functional equations $f(xy)=f(x)+f(y)$, $f(x^n)=nf(x)$. By JOSEPH MILKMAN 505
Transitive systems of linear operators on a Banach space. By BERTRAM YOOD 509
On lacunary Dirichlet series. By I. I. HIRSCHMAN and J. A. JENKINS 512
On the absolute convergence of trigonometrical series. By TAMOTSU TSUCHIKURA and SHIGEKI YANO 517
On double Fourier series. By G. I. SUNOUCHI 522
On the strong summability of Fourier series. By G. I. SUNOUCHI 526
On theorems of M. Riesz and Zygmund. By A. P. CALDERÓN .. 533
Summability and analytic continuation. By V. F. COWLING .. 536
A Bessel function expansion. By MORRIS KLINE 543
On a simple property of the derivatives of Legendre's polynomials. By EMIL GROSSWALD 553
Applications of a lemma of Fejér to typically-real functions. By M. S. ROBERTSON 555
On conformal mapping of nearly circular regions. By S. E. WARSCHAWSKI 562

Articles for insertion in the Proceedings should be typewritten, double spaced, and the author should keep a complete copy. Papers in algebra and theory of numbers should be sent to NATHAN JACOBSON, Leet Oliver Memorial Hall, Yale University, New Haven 11, Conn.; in analysis to A. C. SCHAEFFER, North Hall, University of Wisconsin, Madison 6, Wis.; in other fields to G. A. HEDLUND, Leet Oliver Memorial Hall, Yale University, New Haven 11, Conn. All other communications to the editors should be addressed to the Managing Editor, NATHAN JACOBSON, Leet Oliver Memorial Hall, Yale University, New Haven 11, Conn.

[redacted] BANTA PUBLISHING COMPANY, MENASHA, WISCONSIN